

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2003, 20:50:19 ; Search time 30.0888 Seconds
(without alignments)
2001.720 Million cell updates/sec

Title: US-10-004-219B-9

Perfect score: 2493

Sequence: 1 YNLICYFTWAQYRGLGSF.....QQHQAGLVFTSCNCNWP 452

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2089.5	83.8	455	22	AA62541 Disease treatment
2	2089.5	83.8	476	22	AA62543 Disease treatment
3	1662.5	66.7	368	23	ABB7580 Novel human protei
4	1387	55.6	377	22	AA62545 Disease treatment
5	1387	55.6	398	22	AA62544 Disease treatment
6	1301.5	52.2	466	18	AAW08584 Human 50 kDa chiti
7	1301.5	52.2	466	19	AAW40259 Human chitinase pr
8	1301.5	52.2	466	20	AAW42425 MO-218 clone of hu
9	1301.5	52.2	466	22	AAE00432 Human chitinase pr
10	1301.5	52.2	466	23	ABB76291 Human chitinase.

11	1295.5	52.0	466	19	AAW40260 Human chitinase pr
12	1295.5	52.0	466	20	AAW42426 MO-13B clone of hu
13	1295.5	52.0	466	22	AAE00433 Human chitinase pr
14	1295.5	52.0	466	23	ABB76292 Human chitinase.
15	1294.5	51.9	466	18	AAW31498 Human chitotrioid
16	1279.5	51.3	464	18	AAW31497 Human chitotrioid
17	1150.5	46.1	373	19	AAW40261 Human chitinase pr
18	1150.5	46.1	373	20	AAW42427 Clone of the C-ter
19	1150.5	46.1	373	23	ABB76293 Human chitinase tr
20	1148	46.0	373	19	AAW40262 Human chitinase pr
21	1148	46.0	373	20	AAW42428 Chitinase amino ac
22	1148	46.0	373	23	ABB76294 Human chitinase an
23	1146	46.0	387	18	AAW08585 Human 39 kDa chiti
24	1051	42.2	718	16	AAW73992 Murine oviduct spe
25	1042.5	41.8	668	16	AAW73993 Hamster oviduct sp
26	1010	40.5	537	16	AAW73991 Bovine oviduct spe
27	1008	40.4	696	22	ABBS8615 Drosophila melanog
28	962.5	38.6	383	21	ABBS8615 Gp38k protein sequ
29	956.5	38.4	383	18	AAW26751 Bovine whey protei
30	956.5	38.4	421	19	AAW53122 Amino acid sequenc
31	948.5	38.0	385	19	AAW53123 Amino acid sequenc
32	948.5	38.0	385	19	AAW47033 Human cartilage gp
33	948.5	38.0	416	19	AAW47034 Human cartilage gp
34	948.5	38.0	421	19	AAW53121 Amino acid sequenc
35	948.5	38.0	423	18	AAW35930 Human cartilage g1
36	879.5	35.3	460	22	ABB64366 Drosophila melanog
37	863.5	34.6	554	18	AAW01824 Manduca sexta larv
38	863.5	34.6	554	21	AAW07183 Manduca sexta gut
39	851	34.1	4498	22	ABBS8595 Drosophila melanog
40	782	31.4	595	22	ABBS8595 Polypeptide involv
41	775	31.1	321	9	AAW81342 Human digestive sy
42	772.5	31.0	203	22	AAW92862 House dust mite (D
43	729.5	29.3	536	21	AAW52525 Der HMW-map polyte
44	729.5	29.3	536	23	AAU96329 House dust mite (D
45	729.5	29.3	555	21	AAW52523 House dust mite (D

ALIGNMENTS

RESULT 1

AA62541

ID AA62541 standard; protein; 455 AA.

XX AA62541;

XX AC

DT 24-AUG-2001 (first entry)

DE Disease treatment related protein SEQ ID NO: 1.

DE Disease treatment; infection; chronic occlusive pulmonary disease;

XX bronchial asthma.

OS Homo sapiens.

XX WO200136633-A1.

PD 25-MAY-2001.

PF 14-NOV-2000; 2000WO-JF08015.

PR 15-NOV-1999; 99JP-0324467.

XX (TAKE) TAKEDA CHEM IND LTD.

XX Nakanishi A, Morita S;

XX WPI; 2001-397791/42.

PT New proteins, peptides and DNA for treatment of bronchial asthma,

PT chronic occlusive lung disease and infectious disease -

PS Claim 1; Page 97-99; 114pp; Japanese.

xx CC The present invention provides the sequence of a protein which can be
CC used in the treatment and prevention of infectious diseases. Inhibitors
CC of the protein can be used to treat bronchial asthma and chronic
CC occlusive pulmonary disease. The present sequence is a protein described
CC in the exemplification of the invention.

SQ Sequence 455 AA; 83.8%; Score 2089.5; DB 22; Length 455;
 Query Match 81.3%; Pred. No. 1.6e-174;
 Best Local Similarity. 40; Mismatches 42; Indels 3; Gaps 1;
 Matches 369; Conservative

Qy	1	YNLICFTTNAQYRPGLSGKPDINDPCLCTHLIYAFAGMNNNEITTIENWDTLYKAFN	60
Db	1	YQLTCTFTTNAQYRPGLSGREPMONIDPCLCTHLIYAFAGRQNNNEITTIENWDTLYQAFN	60
Qy	61	DLKRNKSLKTLIAIGWNFGTAPFTTWTSTSNRQTFITSVKLFRLQYGFGLDLDWEY	120
Db	61	GLXNKSQSLKTLIAIGWNFGTAPFTTANVSTPNRQTFITSVKLFRLQYGFGLDLDWEY	120
Qy	121	PGSRGSPQDKHFTVLVVKEMREAFQBEAJESNRPRLMVTAAVAGGINSIQAGEIPELS	180
Db	121	PGSRGSPQDKHFTVLVQSMREAFQEAQINKPRLMVTAAVAGGINSIQSGYEIPQLS	180
Qy	181	KYLDFTHVMTYDLHSGWEGVTGENSPLYKYPTETGSNAYLVNDVYVMNYKNGCAPAEKLI	240
Db	181	QYLDYITHVMTYDLHSGWEGVTGENSPLYKYPTDTSNAYLVNDVYVMNYKNGCAPAEKLI	240
Qy	241	VGFPYEGHTFILERPDSNGIGAPTSGDGPAGAYTROAGFWAYYEICTFLRSAGATEVMDAS	300
Db	241	VGFPYTGHNFILSNPNTGIGAPTSAGPAGPYAKESGIWAYYEICTFLKNGATQGMADP	300
Qy	301	QEVPIYAYKANEWLYONIISFSVKAOWLKONNFGGAMIWAIDLDDFTGSCFDGCKEPLTS	360
Db	301	QEVPIYAYQGNVWYGYDNIKSFDIKAOWLKHNFEGGAMVWAIDLDDFTGTFCNQGKFFLIS	360
Qy	361	TLNKALIGISTEGCTAPDVPSPEVTPPP---GSGSGSGSGSGSGSGSGGFCADKADGLYPVA	417
Db	361	TLKALGLQASCTAPAOPIETIAPSGSGNGSGSGSGSGSGSGGFCVAVRANGLYPVVA	420
Qy	418	DDRNAFWQINGITYOQHCOAGLVFDTSCNCCNW	451
Db	421	NNRNAFWHCVNGVTYOONCOAGLVFDTSCDCCNW	454

RESULT 2
AAG62543
ID AAG62543 standard: protein: 476 AA.

AC: AAG62543:

DT · 24-AUG-2001 (first entry)

DE Disease treatment related protein SEQ ID NO: 5.

Disease treatment; infection; chronic occlusive pulmonary disease; bronchial asthma.

OS Homo sapiens.

XX. WO200136633-A1.
PN

25-MAY-2001.

14-NOV-2000; 2000WO-JP08015.

PR 15-NOV-1999; 99JP-0324467.

PA (TAKE) TAKEDA CHEM IND LTD.

PI Nakanishi A, Morita S;

15

WPI; 2001-397791/42.

New proteins, peptides and DNA for treatment of bronchial asthma, chronic occlusive lung disease and infectious disease -

Example 5; Page 101-103; 114pp; Japanese.

The present invention provides the sequence of a protein which can be used in the treatment and prevention of infectious diseases. Inhibitors of the protein can be used to treat bronchial asthma and chronic occlusive pulmonary disease. The present sequence is a protein described in the exemplification of the invention.

Seq	Sequence	476 AA;
	Query Match	83.8%; Score 2089.5; DB 22; Length 476;
	Best Local Similarity	81.3%; Pred. No. 1.7e-174;
	Matches 369; Conservative	40; Mismatches 42; Indels 3; Gaps 1
QY	1 YNLICYFTNWAQYRPGLSGPKPDDINPCLCTHLIIYAFAGQNNNEITTIENWNVTLKYAFN 60	: : : : :
Db	22 YQLTCYFTNWAQYRPGLSGRFWPONIDPCLCTHLIIYAFAGQNNNEITTIENWNVTLQAFN 81	: : : : :
QY	61 DLKNRNSKLTLLAIGCWNEGTAPFTTWSTSONRQTFITSVIKFLQYFCGLDLDWEY 120	: : : : :
Db	82 GLKNKNSQLKTLTLAIGCWNEGTAPFTTWSTPENRQTFITSVIKFLQYFEGLDLDWEY 141	: : : : :
QY	121 PGRSGSPQDDKHLFTVLVQEMRAFEQEAETESNRPRLMVTAAVAGGISNIQAGYEIPELS 180	: : : : :
Db	142 PGRSGSPQDDKHLFTVLVQEMRAFEQEAQINKPRLMVTAAVAGGISNIQSGYEIPQLS 201	: : : : :
QY	181 KYLDFIHVMTYDLHGSWEGYTGENSEPLIYKYPTFTGSNAYLNVDVYVMYKNGGAPAEKLI 240	: : : : :
Db	202 QYLDVIHVMTYDLHGSWEGYTGENSEPLIYKYPTDTGSNAYLNVDVYVMYKNGGAPAEKLI 261	: : : : :
QY	241 VGPPEYGHFTFLRNPSSNGIGAPTSGDGPAGATROAGFWAYYEICTFLRSGATEWVDAS 300	: : : : :
Db	262 VGFTYGHNFILSNPSTGTIGAPTSAGPAPYAKESGIWAYYEICTFLKNGATQGWDP 321	: : : : :
QY	301 QEVPAVKANEWLGYDNIKSPFSVKAOWLKONFFGGAMIWAIDLDDFTGSCDQGFPLTS 360	: : : : :
Db	322 QEVPAVQGNVWVGVDNIKSPFDIAQWLKRNKKGAMWAIDLDDFTGFCNQGFPLIS 381	: : : : :
QY	361 TLNKAIGISTEGCTAPDVPSEPVTTTP- --GSGSGSGSGSGSGSGGFCADKADGLYPVA 417	: : : : :
Db	382 TLKXALGLQSASCTAPAPQIEPITAPSGSGSGSGSGSGSGGFCVAVRANGLYPVA 441	: : : : :
QY	418 DDRNAPWQCINGITYQQHQAGLVFDTSCNCCNW 451	: : : : :
Db	442 NNRNAPVHCNVGVTYQONCOAGLVFDTSCDCNW 475	: : : : :

RESULT 3
ABB97580

ABB97380
ID ABB97580 standard: protein: 368 AA.

AA
AC

DT 27-JUN-2002 (first entry)

Novel human protein SEQ ID NO: 848.

Human; antinaeamic; vulnery; antiinflammatory; immunomodulator;
antiinfertility; cerebroprotective; cytostatic; rheumatic; gene therapy;
neuroprotective; antiparkinsonian; protein therapy; EST;
expressed sequence tag.

XX Homo sapiens.

AX
PN
WO200222660-A2.

21-MAR-2002.

PF 10-SEP-2001; 2001WO-US26015.

PR 11-SEP-2000; 2000US-0659671.

XX (HYSE-) HYSEQ INC.

PA Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;

XX Xue AJ, Yang Y, Wehrman T, Drmanac RT;

PI WPI; 2002-292408/33.

XX N-PSDB; AEN32766.

DR An isolated polynucleotide for treating diseases associated with its

XX encoded polypeptide such as cancer and multiple sclerosis -

PT Claim 20; SEQ ID NO 848; 509pp; English.

PS The present invention provides the protein and coding sequences of 444

XX novel human proteins. These were isolated from expressed sequences tags

CC (ESTs). They can be used to stimulate cell growth, to regulate

CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth

CC e.g. in burn treatment, to regulate the immune system e.g. to treat

CC multiple sclerosis, to regulate activin or inhibit e.g. to treat

CC infertility, to regulate haemostasis or thrombolysis e.g. to treat

CC stroke and cancer, to screen for drugs, to treat inflammatory conditions

CC e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.

CC Parkinson's disease. The present sequence is a protein of the invention.

XX SQ Sequence 368 AA;

Query Match

Best Local Similarity

Matches 294; Conservative 35; Mismatches 35; Indels 3; Gaps 1;

38 MYSTQNRQTFITSVIKFLRQYFGDGLDWEYDPSGRGSPPODKHLFTVLVKEAREAFQ 147

DB 1 MYSTPENRQTFITSVIKFLRQYFGDGLDWEYDPSGRGSPPODKHLFTVLVKEAREAFQ 60

QY 148 EAIENRRLMTAAVAGGISNIQAGYEIPELSKYLDFLHVMTYDLHGSGWEGYTGNSPL 207

DB 61 EAKQINKPLMTAAVAGGISNIQAGYEIPELSKYLDFLHVMTYDLHGSGWEGYTGNSPL 120

QY 208 YKYPTETGSNAYLVNDYVNMVYKNGAPAEKLIIVGPEYGHFTFLRNPDSNGIGAPTSGD 267

DB 121 YKYPTDTGSNAYLVNDYVNMVYKNGAPAEKLIIVGPEYGHFTFLRNPDSNGIGAPTSGA 180

QY 268 GPAGAYTROAGFWAYEICTFLRSAGTEVWDASQEVPPYAKANEMLYDNIKSFVSKAOW 327

DB 181 GPAGPYAKESGIWAYEICTFLKNGATOGWDAPQEVPPYQGVNMVYDNIKSFIDIKAO 240

QY 328 LKONNFGGAMIWALDDFTGSCDQKPLTSTLNKALGISTEGCTAPDVPSEPVTTTP 387

DB 241 LKHNFPGGAWWALDDFTGFCNGKPLSTLTKALGLOSACCTAPQPIETIAP 300

QY 388 ---GSGSGGSGSGSGSGGFCADKADGLYPVADDRNFAWOCINGITYOQHCAQLVFT 444

DB 301 SCSGNGSGSGSGSGSGSGGFCVAVRANGLYPVANNENAFMHCVNGVITYQNCQAGLVFT 360

QY 445 SCNCCNW 451

DB 361 SCDCCNW 367

RESULT 4

AAG62545

ID AAG62545 standard; protein; 377 AA.

XX AAG62545;

XX AAG62545;

DT 24-AUG-2001 (first entry)

XX Disease treatment related protein SEQ ID NO: 18.

XX

KW Disease treatment; infection; chronic occlusive pulmonary disease;

XX bronchial asthma.

OS Mus sp.

XX WO200136633-A1.

XX 25-MAY-2001.

XX 14-NOV-2000; 2000WO-JP08015.

XX 15-NOV-1999; 99JP-0324467.

XX (TAKE) TAKEDA CHEM IND LTD.

XX Nakanishi A, Morita S;

XX WPI; 2001-397791/42.

XX New proteins, peptides and DNA for treatment of bronchial asthma,

XX chronic occlusive lung disease and infectious disease -

XX Claim 18; Page 110-112; 114pp; Japanese.

XX The present invention provides the sequence of a protein which can be

XX used in the treatment and prevention of infectious diseases. Inhibitors

XX of the protein can be used to treat bronchial asthma and chronic

XX occlusive pulmonary disease. The present sequence is a protein described

XX in the exemplification of the invention.

SQ Sequence 377 AA;

Query Match

Best Local Similarity

Matches 249; Conservative 55; Mismatches 72; Indels 0; Gaps 0;

QY 1 YNLICYFTNWAQYRFLGSGFKPDDINPCLCTHLIYAFAGMQNNEITTIENNDVTLYKAFN 60

DB 1 YQLMCIYTSWAKDRPIEGSKFPGNIDPCLCTHLIYAFAGMQNNEITYTHEQDLRDYEALN 60

QY 61 DLKRNRSKLTLAIGGNWFTAPTITWVSTSONQOTITSVIKFLRQYGFPLGLDWEY 120

DB 61 GLKDKNTLKTLLAIGGNWFTAPTITWVSTSONQOTITSVIKFLRQYGFPLGLDWEY 120

QY 121 PGRSGSPQDDKHLFTVLVKEAREAFQEAIESNRRLMTAAVAGGISNIQAGYEIPELS 180

DB 121 PGRSGSPQDDKHLFTVLVKEAREAFQEAIESNRRLMTAAVAGGISNIQAGYEIPELS 180

QY 181 KYLDFIHWMTYDLHGSGWEGYTGNSPLYKYPTETGSNAYLVNDYVNMVYKNGAPAEKLI 240

DB 181 QSLDYIQWMTYDLHGSGWEGYTGNSPLYKYPTETGSNAYLVNDYVNMVYKNGAPAEKLI 240

QY 241 VGFPYGHFTFLRNPDSNGIGAPTSGDGPAGAYTROAGFWAYEICTFLRSAGTEVWDAS 300

DB 241 VGFPYGHFTFLRNPDSNGIGAPTSGDGPAGAYTROAGFWAYEICTFLRSAGTEVWDAS 300

QY 301 QEVPPYAKANEMLYDNIKSFVSKAQLKONNFGGAMIWALDDFTGSCDQKPLTSTLNK 360

DB 301 QEVPPYAKANEMLYDNIKSFVSKAQLKONNFGGAMIWALDDFTGSCDQKPLTSTLNK 360

QY 361 TLNKALGISTEGCTAP 376

DB 361 TLNKALGISTEGCTAP 376

RESULT 5

AAG62544

ID AAG62544 standard; protein; 398 AA.

XX AAG62544;

XX AAG62544;

DT 24-AUG-2001 (first entry)

XX

DE Disease treatment related protein SEQ ID NO: 17.
XX Disease treatment; infection; chronic occlusive pulmonary disease;
KW bronchial asthma.
XX Mus sp.
XX WO200136633-A1.
XX 25-MAY-2001.
XX 14-NOV-2000; 2000WO-JP08015.
XX 15-NOV-1999; 99JP-0324467.
XX (TAKE) TAKEDA CHEM IND LTD.
XX Nakanishi A, Morita S;
XX WPI; 2001-397791/42.
XX New proteins, peptides and DNA for treatment of bronchial asthma,
PT -Chronic occlusive lung disease and infectious disease -
XX Disclosure; Page 108-110; 114pp; Japanese.
XX The present invention provides the sequence of a protein which can be
CC used in the treatment and prevention of infectious diseases. Inhibitors
CC of the protein can be used to treat bronchial asthma and chronic
CC occlusive pulmonary disease. The present sequence is a protein described
CC in the exemplification of the invention.
XX Sequence 398 AA;
SQ
Query Match 55.6%; Score 1387; DB 22; Length 398;
Best Local Similarity 66.2%; Pred. No. 5.4e-113;
Matches 249; Conservative 55; Mismatches 72; Indels 0; Gaps 0;
QY 1 YNLICVFTWAQVRPGLSGFKPDDINPCLCTHLIYAFAGQNNNEITTIEMNDVTLKAFN 60
DB 22 YQLMCYFTSWAKDRPIEGSGFKPNIDPCLCTHLIYAFAGQNNNEITTIEMNDVTLKAFN 81
QY 61 DLKRNKSLKTLAIGGWNFGTAPFTTWTSONRQTFTITSVIFLRQYCFDGLDLDWEY 120
DB 82 GLKDKNTELKTLAIGGWNFGTAPFTTWTSONRQTFTITSVIFLRQYCFDGLDLDWEY 141
QY 121 PGRSGPPQDKHLFTVLVXEMREAFQEAIESNRPRLMTAAVAGISNIQAGYIPELS 180
DB 142 PGRSGPPQDKHLFTVLVXEMREAFQEAIESNRPRLMTAAVAGISNIQAGYIPELS 201
QY 181 KYLDFTVMTYDLHGSWEGYTGNSPLYKYPTFTGNSNAYLVNDVYMNKNGAPAEKLI 240
DB 202 QSLDYIQVMTYDLHDPKDGTYGNSPLYKSPDYDYGKSAADLVNDSIISYWKDHGAASEKLI 261
QY 241 VGPPEYCHTFLRNPNDNGIGTAPSGDGAGATROAGFWAYYEICTFLRSGATEVWDS 300
DB 262 VGPPEYCHTFLRNPNDNGIGTAPSGDGAGATROAGFWAYYEICTFLRSGATEVWDS 321
QY 301 QEVPAKANEMLYGYNISFVSVAQMLKONNFGGAMIWAILDDPTGFCQKGFPLTS 360
DB 322 QEVPAKANEMLYGYNISFVSVAQMLKONNFGGAMIWAILDDPTGFCQKGFPLTS 381
QY 361 TLNKGALISTEGCTAP 376
DB 382 TLNKGALISTEGCTAP 397

RESULT 6
AAW08584
ID AAW08584 standard; Protein; 466 AA.
XX AAW08584;
AC AAW08584;
XX

DT 24-MAR-1997 (first entry)
XX Human 50 kDa chitinase.
KW Chitinase; chitosinase; chitin; infectious disease;
KW gene therapy; vaccine; diagnosis; lysosomal lipodosis;
KW Gaucher disease; leishmaniasis; sarcoidosis; multiple sclerosis;
KW X-linked adrenoleukodystrophy; drug delivery; cosmetics; food.
XX Homo sapiens.
XX Key Location/Qualifiers
FH Peptide 1..21
FT /label= Sig_peptide
FT Domain 178..198
FT /label= Catalytic_domain
XX WO9640940-A2.
XX 19-DEC-1996.
XX 06-JUN-1996; 96WO-NL00225.
XX 07-JUN-1995; 95US-0486839.
XX (UNAM) UNIV AMSTERDAM.
XX Aerts JMG;
XX WPI; 1997-118698/11.
XX N-PSDB; AAT50833.
XX New human chitinase - used to treat or prevent infection by
PT chitin-contg. pathogens, in diagnosis and as additives to cosmetics,
PT foods, implants etc.
XX Claim 1; Page 40-42; 58pp; English.
XX Human chitinases of 50 kDa (AAW08584) and 39 kDa (AAW08585) are stable
CC to many proteases, active at pH 3-8 and up to 50 deg, and stable in
CC the circulation. They are the result of alternative splicing of
CC RNA, and can be produced on a large-scale in transformed host cells
CC using cDNA clones (see also AAT50833-34) obt'd. from a human macrophage
CC library. The chitinases, optionally expressed from a gene therapy
CC vector, are used to treat or prevent infection by chitin-contg.
CC pathogens (e.g. fungi, protozoa, helminths) and more generally to
CC degrade chitin. They are also useful in controlled-release drug
CC delivery, in cosmetics, foods and dental products, for antibody
CC prodn. and for diagnosis of diseases associated with elevated
CC chitinase levels.
XX Sequence 466 AA;
SQ
Query Match 52.2%; Score 1301.5; DB 18; Length 466;
Best Local Similarity 51.5%; Pred. No. 2.1e-105;
Matches 235; Conservative 71; Mismatches 129; Indels 21; Gaps 3;
QY 3 LICYTNAQVRPGLSGFKPDDINPCLCTHLIYAFAGQNNNEITTIEMNDVTLKAFNDL 62
DB 24 LVCYFTNAQVRPGLSGFKPDDINPCLCTHLIYAFAGQNNNEITTIEMNDVTLKAFNDL 83
QY 63 KURNKSLKTLAIGGWNFGTAPFTTWTSONRQTFTITSVIFLRQYCFDGLDLDWEY 122
DB 84 KURNKSLKTLAIGGWNFGTAPFTTWTSONRQTFTITSVIFLRQYCFDGLDLDWEY 143
QY 123 SRGSPQDKHLFTVLVXEMREAFQEAIESNRPRLMTAAVAGISNIQAGYIPELSKY 182
DB 144 SQGSPAVDKERFTLLVQDLANAFQEAQTSQKRLLSAAVPAQOTYVDAGYEVKIAQN 203
QY 183 LDFHVMYDLHGSWEGYTGNSPLYKYPTFTGNSNAYLVNDVYMNKNGAPAEKLI 242
DB 204 LDFHVMYDLHGSWEGYTGNSPLYKYPTFTGNSNAYLVNDVYMNKNGAPAEKLI 263


```
Query Match      .52.2%; Score 1301.5; DB 22; Length 466;
Best Local Similarity 51.5%; Pred. No. 2.1e-105;
Matches 235; Conservative 71; Mismatches 129; Indels 21; Gaps 3;

QY 3 LICVFTNWAQYRPGLSFKPDDINPCLCTHLYIYAPAGQNNNEITTIENVDVLYKAFNDL 62
DB 24 LVCYFTNWAQYRQGEAREFLPKDLPCLCTHLYIYAPAGTNHQLSTTEWNETLYQEFNGL 83

QY 63 KNRNSKLTLLAIGGNWFGTAPFTTWSTSONROTFTSVIKFLROYGFDGLDLDWEYPG 122
DB 84 KKNMPKLTLLAIGGNWFGTQKFTDMVATANNROTTFVNSAIRFLKYSFDGLDLDWEYPG 143

QY 123 SRGSPPODKHLFTVLVKEMREAFQEAEIESNRPRLMVTAAVAGGISNIQAQYEIPELSKY 182
DB 144 SQGSPAVDKERFTTLVQDLANAFQEAQTSKGERLLLSAAVPAGQTYDAGYEVDKIAQN 203

QY 183 LDFIHVMYDLHGSWEGYTGNSPLYKYPTETGNSNAYLNVDVVMYKNGNAPAKLIVG 242
DB 204 LDFVNLMAFYDFHGSWEKVTGHNSPLYKQESGAASLNVDAAVQOQLQKGTTPASKLILG 263

QY 243 PPEYCHTILRNPDSNGIGAPTSGDPGAGAYTRQAGFWAYYEICTFLRSAGTEVWDASOE 302
DB 264 MPTYGRSFTLASSSDTRVRGAPATSGTGPFTKGGMLAYYEVCSW--KGATKQRIQDQK 321

QY 303 VPYAYKANWELGYDNIKSFSVKAQWLKQNNFGGAMIWALDLDFTGSCDQKFPILTSTL 362
DB 322 VPYIFRDQWVGFDVSEFKTKVYLKQKGLGGAMVWALDLDLDFAGFSCNQGRYPLIQT 381

QY 363 NKALGISTEGCTAPDV-----PSEPVTTPPGSGGGSGSGSGGFCADKADGLYP 415
DB 382 RQELSLPLPSGTPELEVPKQPSEP-----EHGSPSPQDTFCQKADGLYP 429

QY 416 VADDRNAPWQINGITYOQHCQAGLVFDTSCNCCNW 451
DB 430 NPRESSSFYSCAAGRLFOQSCPTGLVFSNSCKCCTW 465

RESULT 10
ABB76291
ID ABB76291 standard; Protein; 466 AA.
AC ABB76291;
XX
XX
XX 12-AUG-2002 (first entry)
XX
XX Human chitinase.
XX
XX Chitinase; enzyme; human; fungicide; antifungal; infection;
KW candidiasis; aspergillosis; coccidioidomycosis; blastomycosis;
KW paracoccidioidomycosis; histoplasmosis; cryptococcosis;
KW chromoblastomycosis; sporotrichosis; mucormycosis; dermatophytosis;
KW Pneumocystis.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FT Peptide 1..21
FT Protein /label= Signal_peptide
FT Protein /label= Mature_protein
XX
XX US6372212-B1.
XX
XX 16-APR-2002.
XX
XX 16-JUN-1997; 97US-0877599.
XX
XX 14-JUN-1996; 96US-0663618.
XX
XX (ICOS-) ICOS CORP.
XX
XX Gray PW;
```

```
XX WPI; 2002-442449/47.
DR N-PSDB; ABL57380.
XX
PT Co-administering chitinase to improve the effectiveness of fungicidal
PT drugs e.g. amphotericin B or itraconazole, useful for treating fungal
PT infections e.g. candidiasis, coccidioidomycosis and blastomycosis -
XX
PS Claim 1; Column 27-30; 26pp; English.
XX
CC The present sequence is the protein sequence of human chitinase
CC as predicted from isolated cDNA clone MO-218 (see ABL57380). It
CC differs in only 1 amino acid residue from the chitinase sequence
CC (see ABB76292) deduced from a second cDNA clone, having glycine
CC at position 81 of the mature protein. Northern blots showed
CC highest chitinase gene expression in lung and ovary tissues.
CC Expression in lung is consistent with a protective role against
CC pathogenic organisms that contain chitin. The invention provides
CC human chitinase polynucleotides and polypeptides, and materials and
CC methods for the recombinant production of human chitinase products,
CC which are expected to be useful as products for treating fungal
CC infections or for the development of such products. Human
CC chitinase has a synergistic effect on the actions of other
CC fungicides. It can be administered to improve the antifungal
CC activity of a non-chitinase antifungal agent, especially
CC amphotericin B or itraconazole, in the treatment of a fungal
CC infection such as candidiasis, aspergillosis, coccidioidomycosis,
CC blastomycosis, paracoccidioidomycosis, histoplasmosis,
CC cryptococcosis, chromoblastomycosis, sporotrichosis, mucormycosis,
CC dermatophytosis and Pneumocystis infections (all claimed). In
CC particular, the fungal infection involves Candida, Aspergillus
CC and/or Cryptococcus spp., whose growth is not effectively
CC inhibited by contact with human chitinase alone.
XX
SQ Sequence 466 AA;
Query Match      52.2%; Score 1301.5; DB 23; Length 466;
Best Local Similarity 51.5%; Pred. No. 2.1e-105;
Matches 235; Conservative 71; Mismatches 129; Indels 21; Gaps 3;

QY 3 LICVFTNWAQYRPGLSFKPDDINPCLCTHLYIYAPAGQNNNEITTIENVDVLYKAFNDL 62
DB 24 LVCYFTNWAQYRQGEAREFLPKDLPCLCTHLYIYAPAGTNHQLSTTEWNETLYQEFNGL 83

QY 63 KNRNSKLTLLAIGGNWFGTAPFTTWSTSONROTFTSVIKFLROYGFDGLDLDWEYPG 122
DB 84 KKNMPKLTLLAIGGNWFGTQKFTDMVATANNROTTFVNSAIRFLKYSFDGLDLDWEYPG 143

QY 123 SRGSPPODKHLFTVLVKEMREAFQEAEIESNRPRLMVTAAVAGGISNIQAQYEIPELSKY 182
DB 144 SQGSPAVDKERFTTLVQDLANAFQEAQTSKGERLLLSAAVPAGQTYDAGYEVDKIAQN 203

QY 183 LDFIHVMYDLHGSWEGYTGNSPLYKYPTETGNSNAYLNVDVVMYKNGNAPAKLIVG 242
DB 204 LDFVNLMAFYDFHGSWEKVTGHNSPLYKQESGAASLNVDAAVQOQLQKGTTPASKLILG 263

QY 243 PPEYCHTILRNPDSNGIGAPTSGDPGAGAYTRQAGFWAYYEICTFLRSAGTEVWDASOE 302
DB 264 MPTYGRSFTLASSSDTRVRGAPATSGTGPFTKGGMLAYYEVCSW--KGATKQRIQDQK 321

QY 303 VPYAYKANWELGYDNIKSFSVKAQWLKQNNFGGAMIWALDLDFTGSCDQKFPILTSTL 362
DB 322 VPYIFRDQWVGFDVSEFKTKVYLKQKGLGGAMVWALDLDLDFAGFSCNQGRYPLIQT 381

QY 363 NKALGISTEGCTAPDV-----PSEPVTTPPGSGGGSGSGSGGFCADKADGLYP 415
DB 382 RQELSLPLPSGTPELEVPKQPSEP-----EHGSPSPQDTFCQKADGLYP 429

QY 416 VADDRNAPWQINGITYOQHCQAGLVFDTSCNCCNW 451
DB 430 NPRESSSFYSCAAGRLFOQSCPTGLVFSNSCKCCTW 465
```


CC molecules that specifically bind to the chitin-binding domain of human
CC chitinase or that modulate its activity. These compounds are useful for
CC immunization, as well as for purifying chitinase, as well as for
CC detection and quantification of chitinase. Polynucleotide fragments of
CC the invention are useful as a source of probes and primers, and to
CC express the proteins recombinantly. The chitinase fragments, when
CC conjugated to antifungal compounds, are used to treat animals,
CC especially humans, infected with chitin-containing parasites such as
CC fungi. Fungal infection treated include candidiasis, aspergillosis,
CC coccidioidomycosis, blastomycosis, paracoccidioidomycosis,
CC mucormycosis, histoplasmosis, cryptococcosis, chromoblastomycosis,
CC sporotrichosis, and dermatophytoses.
CC Chitin can be degraded by the enzyme chitinase. Use of whole chitinase
CC protein for treating infections, especially fungal infections, is
CC problematic. In view of the increasing incidents of life-threatening
CC fungal infection in e.g. immunocompromised individuals, there exists a
CC need for identifying new compounds for treating fungal infection. The
CC chitin-binding fragments of the present invention provide this need.
XX
XX
SQ Sequence 466 AA;

Query Match 52.0%; Score 1295.5; DB 20; Length 466;
Best Local Similarity 51.3%; Pred. No. 7.2e-105;
Matches 234; Conservative 71; Mismatches 130; Indels 21; Gaps 3;
QY 3 LICFTNWAQVRPGLGSKFPDDINPCLCTHLYAFAGMONEITTIENVDVLYKAFNDL 62
DB 24 LVCYFTNWAQVRQGEARFLPKDLSLCTHLYAFAGMTHQSLSTWDETLYQEFNGL 83
QY 63 KNRNSKLTLAIGGNWFTGAPFTTMTVSTQNRQFTITSVKFLRQYFGDGLDLDWEYPG 122
DB 84 KMNPKLTLAIGGNWFTGAPFTTMTVSTQNRQFTITSVKFLRQYFGDGLDLDWEYPG 143
QY 123 SRGSPQDKHLFTVLVKEMREAFQEAIESNRPRLMTAAVAGGISNIOAGYEIPELSKY 182
DB 144 SQGSPAVDKERFTTLVQDLANAFQGEAQTSKGERLLLSAAVPAGQTYVDAGYEVNDKIAQN 203
QY 183 LDFIHVMYDLHGSWEGVTGNSPLYKYPTEGTSNAYLNVDVYMWYKNGAPASKLIVG 242
DB 204 LDFVNLMAFYDFHGSWEKVTGNSPLYKQESGAASLNVDAAVQWQKQTPASKLILG 263
QY 243 PFEYGHFTILRNPDSNGIGAPTSQDPAQAGYTRQAGFWAYBICITFLRSAGTEVNDASQE 302
DB 264 MPTYGRSFTLASSDTRVGPATGSGTPGFTKEGGLAYYEVCSW--KGATKQRIQDOK 321
QY 303 VPYAKANEWLGYNIKSFVKAQWLKQNNFGAMWALDDFTGSCFDQCKPLITSL 362
DB 322 VFIYFRDNQWGFDDVESFTKVSYLKQKGLGAMVWALDLDLDFAGFCNQGRYPLIOTL 381
QY 363 NKALGISTEGCTAPDV-----PSEVPTTPGSGSGSGSGSGGSCFCAKADGLYP 415
DB 382 ROELSLPVLPSGTPELVKPCQPSPE-----EHGSPGQDTFCCKADGLYP 429
QY 416 VADDRNAPMQCINGITYQHQCOAGLVFTSCNCCNW 451
DB 430 NPERSSFYSCAAGRLFQOQSCPTGLVFSNSCKCTW 465

RESULT 13

AAE00433

XX AAE00433 standard; Protein; 466 AA.

XX AAE00433;

XX AAE00433;

XX AAE00433;

DT 19-JUN-2001 (first entry)

XX Human chitinase protein from clone pMO-13B.

KW Human; antifungal; chitinase; immunoglobulin; Ig; therapy;

KW fungal infection; candidiasis; aspergillosis; coccidioidomycosis;

KW blastomycosis; paracoccidioidomycosis; histoplasmosis; cryptococcosis;

KW chromoblastomycosis; sporotrichosis; mucormycosis; dermatophytoses;

KW clone pMO-13B.

XX Homo sapiens.
OS
FH Key Location/Qualifiers
FT Peptide 1..21
FT Protein /label= Signal_peptide
FT /label= Human_mature_chitinase_protein
FT Domain 418..466
FT /label= Chitin_binding_domain
FT Region 1..373
FT /note= "Region with triacetylchitotriose
FT hydrolysing activity"

XX WO200123430-A2.

XX 05-APR-2001.

XX 28-SEP-2000; 2000WO-US26960.

XX 30-SEP-1999; 99US-0409918.

XX (ICOS-) ICOS CORP.

XX Allison DS, Dietsch GN, Gray PW, Shaw KD, Steiner BH;

XX WPI; 2001-266141/27.

XX N-PSDB; AAD03760.

XX Novel chitinase immunoglobulin fusion product, useful for treating
PT fungal infections and reducing the amount of a non-chitinase antifungal
PT agent needed for the treatment -
XX

XX Claim 1; Page 36-38; 39pp; English.

CC The present invention relates to a chitinase immunoglobulin (Ig) fusion
CC product, comprising a human chitinase fused to at least a portion of an
CC immunoglobulin chain. The fusion product is useful for treating fungal
CC infections (mycoses) such as candidiasis, aspergillosis, blastomycosis,
CC coccidioidomycosis, paracoccidioidomycosis, histoplasmosis, mucormycosis,
CC cryptococcosis, chromoblastomycosis, sporotrichosis and dermatophytoses.
CC The fusion protein is useful for reducing the amount of non-chitinase
CC antifungal agent needed to exert an antifungal activity. The fusion
CC protein is also useful for preparing
CC a medicament for the prophylactic or therapeutic treatment of fungal
CC infections. Chitinase immunoglobulin fusion product has unexpectedly
CC improved serum half-life and formulation properties.
CC The present sequence is human chitinase protein from clone pMO-13B.
CC Chitinase enzyme degrades chitin which is a homopolymer of
CC beta-(1,4)-linked N-acetylglucosamine residues.

XX Sequence 466 AA;

Query Match 52.0%; Score 1295.5; DB 22; Length 466;
Best Local Similarity 51.3%; Pred. No. 7.2e-105;
Matches 234; Conservative 71; Mismatches 130; Indels 21; Gaps 3;

QY 3 LICFTNWAQVRPGLGSKFPDDINPCLCTHLYAFAGMONEITTIENVDVLYKAFNDL 62

DB 24 LVCYFTNWAQVRQGEARFLPKDLSLCTHLYAFAGMTHQSLSTWDETLYQEFNGL 83

QY 63 KNRNSKLTLAIGGNWFTGAPFTTMTVSTQNRQFTITSVKFLRQYFGDGLDLDWEYPG 122

DB 84 KMNPKLTLAIGGNWFTGAPFTTMTVSTQNRQFTITSVKFLRQYFGDGLDLDWEYPG 143

QY 123 SRGSPQDKHLFTVLVKEMREAFQEAIESNRPRLMTAAVAGGISNIOAGYEIPELSKY 182

DB 144 SQGSPAVDKERFTTLVQDLANAFQGEAQTSKGERLLLSAAVPAGQTYVDAGYEVNDKIAQN 203

QY 183 LDFIHVMYDLHGSWEGVTGNSPLYKYPTEGTSNAYLNVDVYMWYKNGAPASKLIVG 242

DB 204 LDFVNLMAFYDFHGSWEKVTGNSPLYKQESGAASLNVDAAVQWQKQTPASKLILG 263

Search completed: June 29, 2003, 20:59:39
Job time : 32.0888 secs

PD	09-OCT-1997.	
XX		
PF	21-MAR-1997; 97WO-US05072.	
XX		
XX	29-MAR-1996; 96US-0014295.	
XX		
PA	(HUMA-) HUMAN GENOME SCI INC.	
PA	(SMIK) SMITHKLINE BEECHAM CORP.	
XX		
PI	Chopra A, Gentz R, Kirkpatrick RB, Ni J, Thotakura NR,	
XX		
XX	WPI; 1997-503041/46.	
DR	N-PSDB; AAT89181.	
XX		
PT	New isolated human chitotriosidase gene - used to develop products	
PT	for the diagnosis and treatment of tissue remodeling disorders, e.g.	
PT	rheumatoid arthritis	
XX		
PS	Disclosure; Page 22-24; 34pp; English.	
XX		
CC	This protein comprises a variant of human chitotriosidase (see	
CC	also AA031497). Its amino acid sequence was deduced from the	
CC	coding sequence of a full-length cDNA clone (see AAT89181). The	
CC	following are claimed: (1) a nucleic acid sequence encoding	
CC	chitotriosidase; (2) a method of diagnosing a tissue remodeling	
CC	disorder related to expression of a mutated chitotriosidase protein	
CC	in a host comprising carrying out nucleic acid amplification; and	
CC	(3) a method of detecting altered expression of a chitotriosidase	
CC	protein in a host comprising contacting a bodily sample with an	
CC	antibody. Chitotriosidase can degrade extracellular matrix	
CC	substrates with a similar carbohydrate structure to chitin. It can	
CC	be used to develop products which can be used in the diagnosis and	
CC	treatment of tissue remodeling disorders such as rheumatoid	
CC	arthritis or atherosclerosis.	
XX		
SQ	Sequence 466 AA;	
Query Match 51.9%; Score 1294.5; DB 18; Length 466;		
Best Local Similarity 51.3%; Pred. No. 8.8e-105;		
Matches 234; Conservative 71; Mismatches 130; Indels 21; Gaps 3;		
QY	3 LICYFTNNAQYRPLGSGFKPDDINPCLCTHLIYAFAGMONNEITWNVDTLYKAPNDL	62
DB		
DB	24 LVCYFTNNAQYRQGEARFLPKDLDPDLSCLTHLIYAFAGMTNHLSTWVNDLTYQEFNGL	83
QY	63 KNRNSKLTLLAIGGNFCTAPFTTWVSTSONROTFTSVIKELRYQYGFGLDLDWEYPG	122
DB		
DB	84 KQNPPLKLTLLAIGGNFXTQKFTDMVATANNRQTFVNSAIRFLRYKSFGLDLDWEYPG	143
QY	123 SRGSPQDKHLFTVLVKEMREAFQEAIESNRPRLMVTAAVAGGISNIOAGYEIPELSKY	182
DB		
DB	144 SQGSPAVDKERTTLLVQDLANAFQEAQTSKXERLLLSNAVPAGQTYVDAGYEVDKIAQN	203
QY	183 LDFIHMTYDLHGSNEGVTGENSPLYKYPTETGSNAYLNVDYVMYKNGKAPAEKLIYG	242
DB		
DB	204 LDFVNLMAFYDFHGSWEKVTGHSNPLYKQESGAAASLNVDAAVQWLQKGTGPKSLILG	263
QY	243 PSEYGHFTILRNPSDNGIGAPTSGDGPGAGAYTRQAGFWAYYEICTFLRSGATEWWDASQE	302
DB		
DB	264 MEYGRSRFTLASSSDTRVGPATGSGTGPFTKEGGLAYEYEVCSW--KGATKQRIQDQK	321
QY	303 VPYAYKANEWLYGDNKGFVSVAQWLKONNFGGAMIWALDLDFTGSCDCQGFPLTSTL	362
DB		
DB	322 VPYIFRDNQWGFDDVESFKTVSVLKQKGLGGMVWALDLDLDFAGFCNQGRIYLIQTL	381
QY	363 NKALGISTEGCTAPDV-----PSEPVTPPGSGSGSGSGSGGFCADKADGLYP	415
DB		
DB	382 RQELSLPLVLPSCGTPLEVPKGPQSEP-----EHGSPSQDTFCQKADGLYP	429
QY	416 VADDRNATFWQINGITYQHQCGAGLVFDTSCNCCNW	451
DB		
DB	430 NPERSSFSYCAAGRLFQGCPTGLVFNNSCKCCTW	465

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2003, 20:57:59 ; Search time 11.7012 seconds
(without alignments)
1136.565 Million cell updates/sec

Title: US-10-004-219B-9
Perfect score: 2493
Sequence: 1 YNLICYFTNWAQYRPLGSGF.....QQHCQAGLVFTSCNCCNWP 452

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1301.5	52.2	466	US-08-486-839-4	Sequence 4, Appli
2	1301.5	52.2	466	US-09-151-011-4	Sequence 4, Appli
3	1301.5	52.2	466	US-09-039-198A-2	Sequence 2, Appli
4	1301.5	52.2	466	US-09-343-623-4	Sequence 4, Appli
5	1301.5	52.2	466	US-08-877-599-2	Sequence 2, Appli
6	1301.5	52.2	466	US-09-267-574-2	Sequence 2, Appli
7	1295.5	52.0	466	US-09-039-198A-4	Sequence 4, Appli
8	1295.5	52.0	466	US-08-877-599-4	Sequence 4, Appli
9	1295.5	52.0	466	US-09-267-574-4	Sequence 4, Appli
10	1150.5	46.1	373	US-09-039-198A-14	Sequence 14, Appl
11	1150.5	46.1	373	US-08-877-599-14	Sequence 14, Appl
12	1150.5	46.1	373	US-09-267-574-14	Sequence 14, Appl
13	1148	46.0	373	US-09-039-198A-15	Sequence 15, Appl
14	1148	46.0	373	US-08-877-599-15	Sequence 15, Appl
15	1148	46.0	373	US-09-267-574-15	Sequence 15, Appl
16	1146	46.0	387	US-08-486-839-6	Sequence 6, Appli
17	1146	46.0	387	US-09-151-011-6	Sequence 6, Appli
18	1146	46.0	387	US-09-343-623-6	Sequence 6, Appli
19	973.5	39.0	383	US-09-459-749D-17	Sequence 17, Appl
20	948.5	38.0	385	US-08-694-915-2	Sequence 2, Appli
21	948.5	38.0	416	US-08-694-915-4	Sequence 4, Appli
22	863.5	34.6	554	US-08-524-051-2	Sequence 2, Appli
23	863.5	34.6	554	US-09-052-778-16	Sequence 16, Appl
24	802	32.2	559	US-09-545-814-14	Sequence 14, Appl
25	802	32.2	583	US-09-545-814-2	Sequence 2, Appli
26	802	32.2	583	US-09-545-814-5	Sequence 5, Appli
27	799	32.0	635	US-09-545-814-32	Sequence 32, Appl

28	729.5	29.3	536	4	US-09-292-225-21	Sequence 21, Appl
29	729.5	29.3	555	4	US-09-292-225-15	Sequence 15, Appl
30	729.5	29.3	555	4	US-09-292-225-18	Sequence 18, Appl
31	722	29.0	490	4	US-09-292-225-41	Sequence 41, Appl
32	722	29.0	509	4	US-09-292-225-35	Sequence 35, Appl
33	722	29.0	509	4	US-09-292-225-38	Sequence 38, Appl
34	667.5	26.8	489	4	US-09-545-814-29	Sequence 29, Appl
35	470.5	18.9	389	1	US-07-939-501A-1	Sequence 1, Appli
36	470.5	18.9	389	4	US-08-448-398-7	Sequence 7, Appli
37	470.5	18.9	423	1	US-07-939-501A-10	Sequence 10, Appl
38	470.5	18.9	423	1	US-07-939-501A-12	Sequence 12, Appl
39	455.5	18.3	424	1	US-08-045-289C-2	Sequence 2, Appli
40	455.5	18.3	424	3	US-08-371-680-2	Sequence 2, Appli
41	455.5	18.3	424	5	PCT-US94-01198-2	Sequence 2, Appli
42	434	17.4	442	3	US-09-052-778-2	Sequence 2, Appli
43	418	16.8	371	2	US-08-591-629-2	Sequence 2, Appli
44	409	16.4	377	2	US-08-591-629-8	Sequence 8, Appli
45	389	15.6	452	3	US-09-052-778-15	Sequence 15, Appl

ALIGNMENTS

RESULT 1
US-08-486-839-4
; Sequence 4, Application US/08486839
; Patent No. 5928928
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: A human chitinase, its recombinant
; TITLE OF INVENTION: production, its use for decomposing chitin, its use
; TITLE OF INVENTION: in therapy or prophylaxis against infection diseases.
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11758
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

SEQUENCE CHARACTERISTICS:
LENGTH: 466 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-08-486-839-4

Query Match 52.2%; Score 1301.5; DB 2; Length 466;
Best Local Similarity 51.5%; Pred. No. 1.4e-108;
Matches 235; Conservative 71; Mismatches 129; Indels 21; Gaps 3;
3 LICYFTNWAQYRPLGSGFKPDDINPCLCTHLYAFAGHONNEITTIWVNDVLYKAFNDL 62
24 LVCYFTNWAQYRQSGARFLPKDLPCLCTHLYAFAGHONNEITTIWVNDVLYKAFNDL 83

[illegible]

RESULT 2
 US-09-151-011-4
 ; Sequence 4, Application US/09151011
 ; Patent No. 6057142
 ; GENERAL INFORMATION:
 ; APPLICANT:
 ; TITLE OF INVENTION: A Human Chitinase, Its Recombinant
 ; TITLE OF INVENTION: Production, Its Use For Decomposing Chitin, Its Use in
 ; TITLE OF INVENTION: Therapy or Prophylaxis Against Infection Diseases.
 ; NUMBER OF SEQUENCES: 16
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Hoffmann & Baron, LLP
 ; STREET: 6900 Jericho Turnpike
 ; CITY: Syosset
 ; STATE: New York
 ; COUNTRY: United States of America
 ; ZIP: 11791
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; SOFTWARE:
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/151.011
 ; FILING DATE: 10 - September - 1998
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Morris, Robert C.
 ; REGISTRATION NUMBER: 42,910
 ; REFERENCE/DOCKET NUMBER: 294-32 DIV
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (516) 822-3550
 ; TELEFAX: (516) 822-3582
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 466 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: unknown
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: protein
 ; HYPOTHETICAL: NO
 ; US-09-151-011-4

RESULT 3
US-09-039-198A-2
/ Sequence 2, Application US/09039198A
/ Patent No. 6200951
/ GENERAL INFORMATION:
/ APPLICANT: Gray, Patrick W.
/ APPLICANT: Tjoelker, Larry W.
/ TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
/ NUMBER OF SEQUENCES: 34
/ CORRESPONDENCE ADDRESSES:
/ ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
/ STREET: 233 South Wacker Drive/6300 Sears Tower
/ CITY: Chicago
/ STATE: Illinois
/ COUNTRY: United States of America
/ ZIP: 60606-6402
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/039,198A
/ FILING DATE:
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Rin-Laures, Li-Hsien
/ REGISTRATION NUMBER: 33,547
/ REFERENCE/DOCKET NUMBER: 27866/34391
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (312) 474-6300
/ TELEFAX: (312) 474-0448
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 466 amino acids
/ TYPE: amino acid

```
;
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-039-198A-2

Query Match      52.2%; Score 1301.5; DB 4; Length 466;
Best Local Similarity 51.5%; Pred. No. 1.4e-108;
Matches 235; Conservative 71; Mismatches 129; Indels 21; Gaps 3;

QY 3 LICYFTNWAQYRPGKSPKDDINPCLCTHLYAFAGMNNIITTIWNDVLYKAFNDL 62
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 24 LVCYFTNWAQYRQGEARFLPKDLDPSLCTHLYAFAGMNHQSLTTEWDETLYQBFNGL 83
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 63 KVRNSKLTLLAIGGNFCTAPFTTWSTQNRQTFITSVIKFLROYGFDGLDWEYPG 122
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 84 KKNPKLKTLLAIGGNFCTQKFTDMVATANNRQTFVNSAIRFLKYSFDGLDWEYPG 143
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 123 SRGSPDQKHLFTVLVKEMREAFQEATESNRPRLMVTAAVAGGISNIOAGYEIPELSKY 182
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 144 SQGSPAVDKERFTTLVQDLANAFQGEAQTSGKERLLLSAAVPAGQTYVDAGYEVDKIAQN 203
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 183 LDFIHVMTYDLHSGWEGYTGENSEPLYKYPTETGNSNAYLNVDVYMNWKNGAPAEKLIYG 242
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 204 LDFVNLMAFYDHGSEWYKVTGHSNPLYKQESGAASLNVDAAVQOQLQKGTPTASKLIIG 263
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 243 FPEYGHFTILRNPSNONGIAPTSGDGPGAGAYTRQAGFWAYYEICTFLRSGATEVMDASOE 302
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 264 MPTYGRSFTLASSSDTRVGAPATGSGTPGFTKEGGLAYYEVCSW--KGATKQRIQDOK 321
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 303 VPYAYKANWELGYDNIKSPVKAQWLKONNFGAMIMWALDLDFTGSCDQCKEPLTSTL 362
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 322 VPYIFRDNQWGFDDVESFKTKVYLKOKGLGAMVWALDLDFFAGFCNCGRYPLIQT 381
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 363 NKALGISTEGCTAPDV-----PSEPVTTPPGSGSGSGSGSGSGSGGFCADKADGLYP 415
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 382 ROELSLPYLPSTGTEPELVKPGQPSPE-----EHGSPSQDFTFCQKADGLYP 429
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 416 VADDRNMFQCCINGITYQOHCQAGLVFDTSCNCNW 451
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 430 NPRESSFSYCAAGRLFQOCSCTGLVFSNSCKCCTW 465
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 4
US-09-343-623-4
; Sequence 4, Application US/09343623
; Patent No. 6303118
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: A human chitinase, its recombinant
; TITLE OF INVENTION: production, its use for decomposing chitin, its use
; TITLE OF INVENTION: in therapy or prophylaxis against infection diseases.
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11758
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/343,623
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,839
; FILING DATE: 07-June-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Baron, Ronald J.
; REGISTRATION NUMBER: 29,281
```

```
;
; REFERENCE/DOCKET NUMBER: 294-26
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 466 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; US-09-343-623-4

Query Match      52.2%; Score 1301.5; DB 4; Length 466;
Best Local Similarity 51.5%; Pred. No. 1.4e-108;
Matches 235; Conservative 71; Mismatches 129; Indels 21; Gaps 3;

QY 3 LICYFTNWAQYRPGKSPKDDINPCLCTHLYAFAGMNNIITTIWNDVLYKAFNDL 62
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 24 LVCYFTNWAQYRQGEARFLPKDLDPSLCTHLYAFAGMNHQSLTTEWDETLYQBFNGL 83
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 63 KVRNSKLTLLAIGGNFCTAPFTTWSTQNRQTFITSVIKFLROYGFDGLDWEYPG 122
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 84 KKNPKLKTLLAIGGNFCTQKFTDMVATANNRQTFVNSAIRFLKYSFDGLDWEYPG 143
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 123 SRGSPDQKHLFTVLVKEMREAFQEATESNRPRLMVTAAVAGGISNIOAGYEIPELSKY 182
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 144 SQGSPAVDKERFTTLVQDLANAFQGEAQTSGKERLLLSAAVPAGQTYVDAGYEVDKIAQN 203
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 183 LDFIHVMTYDLHSGWEGYTGENSEPLYKYPTETGNSNAYLNVDVYMNWKNGAPAEKLIYG 242
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 204 LDFVNLMAFYDHGSEWYKVTGHSNPLYKQESGAASLNVDAAVQOQLQKGTPTASKLIIG 263
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 243 FPEYGHFTILRNPSNONGIAPTSGDGPGAGAYTRQAGFWAYYEICTFLRSGATEVMDASOE 302
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 264 MPTYGRSFTLASSSDTRVGAPATGSGTPGFTKEGGLAYYEVCSW--KGATKQRIQDOK 321
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 303 VPYAYKANWELGYDNIKSPVKAQWLKONNFGAMIMWALDLDFTGSCDQCKEPLTSTL 362
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 322 VPYIFRDNQWGFDDVESFKTKVYLKOKGLGAMVWALDLDFFAGFCNCGRYPLIQT 381
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 363 NKALGISTEGCTAPDV-----PSEPVTTPPGSGSGSGSGSGSGGFCADKADGLYP 415
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 382 ROELSLPYLPSTGTEPELVKPGQPSPE-----EHGSPSQDFTFCQKADGLYP 429
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 416 VADDRNMFQCCINGITYQOHCQAGLVFDTSCNCNW 451
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 430 NPRESSFSYCAAGRLFQOCSCTGLVFSNSCKCCTW 465
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 5
US-08-877-599-2
; Sequence 2, Application US/08877599
; Patent No. 6372212
; GENERAL INFORMATION:
; APPLICANT: Gray, Patrick W.
; TITLE OF INVENTION: Chitinase Materials and Methods
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/877,599
```

[illegible]

CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/663,618

US-09-267-574-4

Query Match 52.0%; Score 1295.5; DB 4; Length 466;
Best Local Similarity 51.3%; Pred. No. 4.7e-108;
Matches 234; Conservative 71; Mismatches 130; Indels 21; Gaps 3;

QY 3 LICYFTNWAQYRPGSGFKPDDINPCLCTHLIYAFAGMONEITTIWINDVLYKAFNDL 62
DB 24 LVCYFTNWAQYRQGEARFLPKDLPSLCTHLIYAFAGMTHQSLSTWENDETLQEFNGL 83
QY 63 KURNKSLKTLAIGGNWFTAPFTTWMVSTSONROTFTITSVKFLROYGDFGLDLWEYPG 122
DB 84 KQNPPLKTLAIGGNWFTQKFTDMVATANNRQTFVNSAIRFLRKYSPDGLDLWEYPG 143
QY 123 SRGSPQDKHLFTVLKEMREAFQEAIESNRPRLMVTAAVAGGISNIQAGYEIPELSKY 182
DB 144 SQSPAVDKERTTLVQDLANAFQQAQTSKGERLLLSAAVPAQTYVDAGYEVVKIAQN 203
QY 183 LDFHVMYDLHGSWEGYTGNSPLYKYPTETGSNAYLNVDYVMYWKNGAPAEKLIIVG 242
DB 204 LDFVNLMAVDFHGSWEKVTGHSNPLYKQESGAASLNVDAAVQQLKQKTPASKLIIG 263
QY 243 PFEYGHFTILRNPDSNGIGAPTSGDGPAGATYRQAGFWAYYEICTFLRSGATEWVDSOE 302
DB 264 MPTYGRSFTLASSSDTRVGPATGSGTPGFTKEGGMLAYEYVCSW--KGATKQRIQDOK 321
QY 303 VPYAYKANWGLVDNIKSFVKAQWLKONNFGGAMIWADLDDFTGSCDQKGFPLTSTL 362
DB 322 VPIYFDNQQWGFDDVESFKTKVSLKQKGLGAMVWALDDDFAGFCNQGRYPLIQT 381
QY 363 NKALGISTEGCTAPDV-----PSEPTVTPPGSGSGSGSGSGSGSGGFCADKADGLYP 415
DB 382 ROELSPLPYLPGCTPELEVPKPGQSEP-----EHGSPRQDFTFCQKADGLYP 429
QY 416 VADDRNFWQCTINGITYQHQQAGLVFDTSCNCCNW 451
DB 430 NPRERSFYSCAAGRLFOQSCPTGLVFNSCRKCTW 465

RESULT 10

US-09-039-198A-14
Sequence 14, Application US/09039198A
Patent No. 6200951
GENERAL INFORMATION:
APPLICANT: Gray, Patrick W.
TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/039,198A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Rin-Laures, Li-Hsien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 27866/34391
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:

LENGTH: 373 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-039-198A-14

Query Match 46.1%; Score 1150.5; DB 4; Length 373;
Best Local Similarity 54.2%; Pred. No. 3.6e-95;
Matches 205; Conservative 65; Mismatches 99; Indels 9; Gaps 2;

QY 3 LICYFTNWAQYRPGSGFKPDDINPCLCTHLIYAFAGMONEITTIWINDVLYKAFNDL 62
DB 3 LVCYFTNWAQYRQGEARFLPKDLPSLCTHLIYAFAGMTHQSLSTWENDETLQEFNGL 62
QY 63 KURNKSLKTLAIGGNWFTAPFTTWMVSTSONROTFTITSVKFLROYGDFGLDLWEYPG 122
DB 63 KQNPPLKTLAIGGNWFTQKFTDMVATANNRQTFVNSAIRFLRKYSPDGLDLWEYPG 122
QY 123 SRGSPQDKHLFTVLKEMREAFQEAIESNRPRLMVTAAVAGGISNIQAGYEIPELSKY 182
DB 123 SQSPAVDKERTTLVQDLANAFQQAQTSKGERLLLSAAVPAQTYVDAGYEVVKIAQN 182
QY 183 LDFHVMYDLHGSWEGYTGNSPLYKYPTETGSNAYLNVDYVMYWKNGAPAEKLIIVG 242
DB 183 LDFVNLMAVDFHGSWEKVTGHSNPLYKQESGAASLNVDAAVQQLKQKTPASKLIIG 242
QY 243 PFEYGHFTILRNPDSNGIGAPTSGDGPAGATYRQAGFWAYYEICTFLRSGATEWVDSOE 302
DB 243 MPTYGRSFTLASSSDTRVGPATGSGTPGFTKEGGMLAYEYVCSW--KGATKQRIQDOK 300
QY 303 VPYAYKANWGLVDNIKSFVKAQWLKONNFGGAMIWADLDDFTGSCDQKGFPLTSTL 362
DB 301 VPIYFDNQQWGFDDVESFKTKVSLKQKGLGAMVWALDDDFAGFCNQGRYPLIQT 360
QY 363 NKALGISTEGCTAPDVPS 380
DB 361 ROELSPLPYLPGCTPELEVPKPGQSEP-----EHGSPRQDFTFCQKADGLYP 429

RESULT 11

US-08-877-599-14
Sequence 14, Application US/08877599
Patent No. 6372212
GENERAL INFORMATION:
APPLICANT: Gray, Patrick W.
TITLE OF INVENTION: Chitinase Materials and Methods
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/877,599
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/663,618
FILING DATE: 14-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Rin-Laures, Li-Hsien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 27866/33994
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300

TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 373 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-877-599-14

Query Match 46.1%; Score 1150.5; DB 4; Length 373;
Best Local Similarity 54.2%; Pred. No. 3.6e-95;
Matches 205; Conservative 65; Mismatches 99; Indels 9; Gaps 2;

QY 3 LICVYTNNAQVRPGSGFKPDDINPCLCTHLIYAFAGMONEITTIENWVDTLYKAFNDL 62
DB 3 LVCYFTNNAQVRQGEARFLPKDLPCLCTHLIYAFAGMTHHOLSTTEWDETLYQEFNGL 62
QY 63 KNRNSKLTLLAIGWNGFTAPFTTMVSTSQNRQFTITSVKIFLRQYFGDGLDLDWEYPG 122
DB 63 KMNPKLTLLAIGWNGFTQKFTDMVATANNRQFTVNSAIRFLKYSFDFGLDLDWEYPG 122
QY 123 SRGSPQDKHLFTVLVKEMREAFQGEAIESNRPRLMTAAVAGGISNIQAGVEIPELSKY 182
DB 123 SQGSPAVDKERFTTLVQDLANAFQGEAQTSGKERLLLSAAVPAQOTYVDAGYEVVDKIAQN 182
QY 183 LDFIHVMYDHLGSGWEGYTGNSPLYKYPTETGNSNAYLNVDVVMYKNNKGAPAEKLI 242
DB 183 LDFVNLMAIDFHSGWEKVTGNSPLYKQESGAASLVNDAVQOVLQKGTTPASKLI 242
QY 243 PPEYGHTEFLRNPDSNGIGAPTSQDGPAGAYTRQAGFWAYVEICTFLRSGATEVWDASQE 302
DB 243 MPTYGRSETLASSSDTRVGAPATGSGTPGPTKEGGLAYEVCWS--KGATKQRIQDOK 300
QY 303 VPIYAYKANWLYGYNIKSFSVKAQWLKQNNFGGAMIWALDDFTGSCDQKFPPLTSTL 362
DB 301 VPIYFRDNQWYGFDDVESFKTKVSYLKQKGLGAMVWALDDDFAGFCNQGRIYPLIOTL 360
QY 363 NKALGISTEGCTAPDVPS 380
DB 361 RQELSL-----PYLPS 371

RESULT 12
US-09-267-574-14
Sequence 14, Application US/09267574
Patent No. 6399571
GENERAL INFORMATION:
APPLICANT: Gray, Patrick W.
TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
FILE REFERENCE: 27866/35407
CURRENT APPLICATION NUMBER: US/09/267,574
CURRENT FILING DATE: 1999-03-12
EARLIER APPLICATION NUMBER: 09/039,198
EARLIER FILING DATE: 1998-03-12
NUMBER OF SEQ ID NOS: 39
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 14
LENGTH: 373
TYPE: PRT
ORGANISM: Homo sapiens
US-09-267-574-14

Query Match 46.1%; Score 1150.5; DB 4; Length 373;
Best Local Similarity 54.2%; Pred. No. 3.6e-95;
Matches 205; Conservative 65; Mismatches 99; Indels 9; Gaps 2;

QY 3 LICVYTNNAQVRPGSGFKPDDINPCLCTHLIYAFAGMONEITTIENWVDTLYKAFNDL 62
DB 3 LVCYFTNNAQVRQGEARFLPKDLPCLCTHLIYAFAGMTHHOLSTTEWDETLYQEFNGL 62

QY 63 KNRNSKLTLLAIGWNGFTAPFTTMVSTSQNRQFTITSVKIFLRQYFGDGLDLDWEYPG 122
DB 63 KMNPKLTLLAIGWNGFTQKFTDMVATANNRQFTVNSAIRFLKYSFDFGLDLDWEYPG 122
QY 123 SRGSPQDKHLFTVLVKEMREAFQGEAIESNRPRLMTAAVAGGISNIQAGVEIPELSKY 182
DB 123 SQGSPAVDKERFTTLVQDLANAFQGEAQTSGKERLLLSAAVPAQOTYVDAGYEVVDKIAQN 182
QY 183 LDFIHVMYDHLGSGWEGYTGNSPLYKYPTETGNSNAYLNVDVVMYKNNKGAPAEKLI 242
DB 183 LDFVNLMAIDFHSGWEKVTGNSPLYKQESGAASLVNDAVQOVLQKGTTPASKLI 242
QY 243 PPEYGHTEFLRNPDSNGIGAPTSQDGPAGAYTRQAGFWAYVEICTFLRSGATEVWDASQE 302
DB 243 MPTYGRSETLASSSDTRVGAPATGSGTPGPTKEGGLAYEVCWS--KGATKQRIQDOK 300
QY 303 VPIYAYKANWLYGYNIKSFSVKAQWLKQNNFGGAMIWALDDFTGSCDQKFPPLTSTL 362
DB 301 VPIYFRDNQWYGFDDVESFKTKVSYLKQKGLGAMVWALDDDFAGFCNQGRIYPLIOTL 360
QY 363 NKALGISTEGCTAPDVPS 380
DB 361 RQELSL-----PYLPS 371

RESULT 13

US-09-039-198A-15
Sequence 15, Application US/09039198A
Patent No. 6200951
GENERAL INFORMATION:
APPLICANT: Gray, Patrick W.
APPLICANT: Tjoelker, Larry W.
TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive/6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/039,198A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Rin-Laures, Li-Hsien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 27866/34391
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 373 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-039-198A-15

Query Match 46.0%; Score 1148; DB 4; Length 373;
Best Local Similarity 55.2%; Pred. No. 6e-95;
Matches 202; Conservative 64; Mismatches 98; Indels 2; Gaps 1;

QY 3 LICVYTNNAQVRPGSGFKPDDINPCLCTHLIYAFAGMONEITTIENWVDTLYKAFNDL 62
DB 3 LVCYFTNNAQVRQGEARFLPKDLPCLCTHLIYAFAGMTHHOLSTTEWDETLYQEFNGL 62

Db 243 MPTYGRSFTLASSSDTRVGPATGSGTGPFTKEGMLAYYEVCSW--KGATKQRIQDQK 300
QY 303 VPYAYKANWLGYNIKSFSVKAQWLKONNFGGAMIWALDLDFTGSCDQKFPPLTSTL 362
Db 301 VPYIPRDNQWGFDDVESFKTKVSYLKQKGLGGANWALDLDLDDFAGSCNQGRYPLIQTL 360
QY 363 NKALGI 368
Db 361 RQELSL 366

Search completed: June 29, 2003, 21:03:28
Job time : 13.7012 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2003, 21:01:56 ; Search time 17.3846 Seconds
(without alignments)
2850.818 Million cell updates/sec

Title: US-10-004-219B-9
Perfect score: 2493
Sequence: 1 YNLICYFTNWAQYRPLGLSF.....QQHQAGLVFTSCNCCNWP 452

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 424699, seqs, 109646833 residues

Total number of hits satisfying chosen parameters: 424699

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US05_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2493	100.0	452	US-10-004-219B-9	Sequence 9, Appli
2	2493	100.0	473	US-10-004-219B-4	Sequence 4, Appli
3	2089.5	83.8	455	US-10-004-219B-14	Sequence 14, Appl
4	2089.5	83.8	476	US-10-004-219B-1	Sequence 1, Appli
5	1301.5	52.2	445	US-10-004-219B-10	Sequence 10, Appl
6	987.5	39.6	383	US-10-097-340-45	Sequence 45, Appl
7	973.5	39.0	383	US-09-459-749D-17	Sequence 17, Appl
8	729.5	29.3	536	US-10-218-743-21	Sequence 21, Appl
9	729.5	29.3	555	US-10-218-743-15	Sequence 15, Appl
10	729.5	29.3	555	US-10-218-743-18	Sequence 18, Appl
11	722	29.0	490	US-10-218-743-41	Sequence 41, Appl
12	722	29.0	509	US-10-218-743-35	Sequence 35, Appl
13	722	29.0	509	US-10-218-743-38	Sequence 38, Appl
14	476	19.1	371	US-09-923-844B-2	Sequence 2, Appli
15	265.5	10.6	376	US-09-748-033-3	Sequence 3, Appli
16	203	8.1	170	US-10-218-743-44	Sequence 44, Appl
17	106	4.3	19	US-10-202-436A-13	Sequence 13, Appl
18	106	4.3	972	US-10-100-049-23	Sequence 23, Appl
19	102	4.1	1385	US-09-738-363-2	Sequence 2, Appli

20	100.5	4.0	400	12	US-10-055-430-1	Sequence 1, Appli
21	98.5	4.0	466	9	US-09-738-626-4763	Sequence 4763, Ap
22	97.5	3.9	1871	9	US-10-184-644-301	Sequence 301, App
23	97.5	3.9	1871	9	US-10-184-634-301	Sequence 301, App
24	96.5	3.9	1434	9	US-10-080-505-9	Sequence 9, Appli
25	96	3.9	237	10	US-09-924-099-9	Sequence 9, Appli
26	96	3.9	243	10	US-09-924-099-10	Sequence 10, Appl
27	95.5	3.8	299	9	US-10-007-521-12	Sequence 12, Appl
28	95.5	3.8	386	9	US-10-050-704-100	Sequence 100, App
29	95.5	3.8	695	10	US-09-945-258-18	Sequence 18, Appl
30	95.5	3.8	1157	9	US-10-184-644-391	Sequence 391, App
31	95.5	3.8	1157	9	US-10-184-634-391	Sequence 391, App
32	95	3.8	309	9	US-10-304-928-10	Sequence 10, Appl
33	95	3.8	661	10	US-09-812-238B-2	Sequence 2, Appli
34	95	3.8	661	10	US-09-862-260A-2	Sequence 2, Appli
35	95	3.8	668	9	US-10-047-539-4	Sequence 4, Appli
36	95	3.8	776	10	US-09-833-435A-4	Sequence 4, Appli
37	94.5	3.8	540	9	US-09-738-626-4450	Sequence 4450, Ap
38	91.5	3.7	425	10	US-09-888-615-92	Sequence 92, Appl
39	91.5	3.7	465	9	US-10-159-487-2	Sequence 2, Appli
40	91.5	3.7	541	9	US-10-071-485-85	Sequence 85, Appl
41	91	3.7	2024	9	US-10-123-155-259	Sequence 259, App
42	91	3.7	2105	9	US-09-428-371-3	Sequence 3, Appli
43	90.5	3.6	434	10	US-09-809-517A-23	Sequence 23, Appl
44	90.5	3.6	434	10	US-09-809-517A-26	Sequence 26, Appl
45	90.5	3.6	440	9	US-10-066-500-106	Sequence 106, App

ALIGNMENTS

RESULT 1
US-10-004-219B-9
; Sequence 9, Application US/10004219B
; Publication No. US20030087414A1
; GENERAL INFORMATION:
; APPLICANT: Macrozyme
; APPLICANT: Aerts, Johannes M.F.G.
; APPLICANT: Boot, Rolf G.
; TITLE OF INVENTION: A mammalian mucinase, its recombinant production, and
; TITLE OF INVENTION: its use in therapy or prophylaxis against diseases in
; TITLE OF INVENTION: which mucus is involved or infection diseases
; FILE REFERENCE: 2183-5136US
; CURRENT APPLICATION NUMBER: US/10/004,219B
; CURRENT FILING DATE: 2001-11-02
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 9
; LENGTH: 452
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1)..(452)
; OTHER INFORMATION: /note="Mouse AMCase"
US-10-004-219B-9

Query Match	100.0%	Score 2493;	DB 9;	Length 452;
Best Local Similarity	100.0%;	Pred. No. 2.5e-205;		
Matches 452;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
y	1	YNLCYFTNWAQYRPLGLSGFKPDDINPCLCTHLIYAFAGMNNETTIEWNDVTLTKAFN	60	
b	1	YNLCYFTNWAQYRPLGLSGFKPDDINPCLCTHLIYAFAGMNNETTIEWNDVTLTKAFN	60	
y	61	DLKNNRSKLTLAIGGWNFGTAPFTTWMVTSQNRQTFTTSVIKFLRQYGFGLDLQWEY	120	
b	61	DLKNNRSKLTLAIGGWNFGTAPFTTWMVTSQNRQTFTTSVIKFLRQYGFGLDLQWEY	120	
y	121	PGSRGSPQDKHLFTVLVKEMREAEQSAIESNRPRLMVTAAVAGGTSNIQAGYEIPELS	180	
b	121	PGSRGSPQDKHLFTVLVKEMREAEQSAIESNRPRLMVTAAVAGGTSNIQAGYEIPELS	180	

QY 181 KYLDFHVTYDLHGSWEGYTGENSEPLYKYPTETGNSNAYLVNDVYNNYKNGGAPAEKLI 240
DB 181 KYLDFHVTYDLHGSWEGYTGENSEPLYKYPTETGNSNAYLVNDVYNNYKNGGAPAEKLI 240
QY 241 VGFPEYGHFTILRNPSDNGIGAPTSQDGPAGAYTRQAGFWAYYEICTFLRSGATEVWDAS 300
DB 241 VGFPEYGHFTILRNPSDNGIGAPTSQDGPAGAYTRQAGFWAYYEICTFLRSGATEVWDAS 300
QY 301 QVVPYAYKANWLGNDYNIKSFSVKAQWLKQNNFGGAMIWAILDDFTGSCDQGFPLTS 360
DB 301 QVVPYAYKANWLGNDYNIKSFSVKAQWLKQNNFGGAMIWAILDDFTGSCDQGFPLTS 360
QY 361 TLNKALGISTEGCTAPDVPSEPTTPPGSGSGSGSGSGSGSGGFCADKADGLYPVADDR 420
DB 361 TLNKALGISTEGCTAPDVPSEPTTPPGSGSGSGSGSGSGSGGFCADKADGLYPVADDR 420
QY 421 NAFWQINGITYQHQHQAAGLVFDTSCNCCNP 452
DB 421 NAFWQINGITYQHQHQAAGLVFDTSCNCCNP 452
RESULT 2
US-10-004-219B-4
; Sequence 4, Application US/10004219B
; Publication No. US20030087414A1
; GENERAL INFORMATION:
; APPLICANT: Macrozyme
; APPLICANT: Aerts, Johannes M.F.G.
; APPLICANT: Boot, Rolf G.
; TITLE OF INVENTION: A mammalian mucinase, its recombinant production, and
; TITLE OF INVENTION: its use in therapy or prophylaxis against diseases in
; TITLE OF INVENTION: which mucus is involved or infection diseases
; FILE REFERENCE: 2183-5136US
; CURRENT APPLICATION NUMBER: US/10/004,219B
; CURRENT FILING DATE: 2001-11-02
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 473
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: mouse AMCase
; OTHER INFORMATION: amino acid sequence deduced from cDNA sequence
US-10-004-219B-4
Query Match 100.0%; Score 2493; DB 9; Length 473;
Best Local Similarity 100.0%; Pred. No. 2.7e-205;
Matches 452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 YNLICVFTNWAQYRPLGSGFKPDDINPCLCTHLIYAFAGMONEITTIENNDVTLYKAFN 60
DB 22 YNLICVFTNWAQYRPLGSGFKPDDINPCLCTHLIYAFAGMONEITTIENNDVTLYKAFN 81
QY 61 DLKRNKSLKTLTAIGGNWFGTAPFTTWVSTSONRQTFTTSVKELROYGFGDLDDWEY 120
DB 82 DLKRNKSLKTLTAIGGNWFGTAPFTTWVSTSONRQTFTTSVKELROYGFGDLDDWEY 141
QY 121 PGSRSGPPQDKHLFTVLVKEMREAFQEAIESNRPLMTAAVAGGISNIQAGYIPELS 180
DB 142 PGSRSGPPQDKHLFTVLVKEMREAFQEAIESNRPLMTAAVAGGISNIQAGYIPELS 201
QY 181 KYLDFHVTYDLHGSWEGYTGENSEPLYKYPTETGNSNAYLVNDVYNNYKNGGAPAEKLI 240
DB 202 KYLDFHVTYDLHGSWEGYTGENSEPLYKYPTETGNSNAYLVNDVYNNYKNGGAPAEKLI 261
QY 241 VGFPEYGHFTILRNPSDNGIGAPTSQDGPAGAYTRQAGFWAYYEICTFLRSGATEVWDAS 300
DB 262 VGFPEYGHFTILRNPSDNGIGAPTSQDGPAGAYTRQAGFWAYYEICTFLRSGATEVWDAS 321
QY 301 QVVPYAYKANWLGNDYNIKSFSVKAQWLKQNNFGGAMIWAILDDFTGSCDQGFPLTS 360
DB 322 QVVPYAYKANWLGNDYNIKSFSVKAQWLKQNNFGGAMIWAILDDFTGSCDQGFPLTS 381

QY 361 TLNKALGISTEGCTAPDVPSEPTTPPGSGSGSGSGSGSGGFCADKADGLYPVADDR 420
DB 382 TLNKALGISTEGCTAPDVPSEPTTPPGSGSGSGSGSGSGGFCADKADGLYPVADDR 441
QY 421 NAFWQINGITYQHQHQAAGLVFDTSCNCCNP 452
DB 442 NAFWQINGITYQHQHQAAGLVFDTSCNCCNP 473
RESULT 3
US-10-004-219B-14
; Sequence 14, Application US/10004219B
; Publication No. US20030087414A1
; GENERAL INFORMATION:
; APPLICANT: Macrozyme
; APPLICANT: Aerts, Johannes M.F.G.
; APPLICANT: Boot, Rolf G.
; TITLE OF INVENTION: A mammalian mucinase, its recombinant production, and
; TITLE OF INVENTION: its use in therapy or prophylaxis against diseases in
; TITLE OF INVENTION: which mucus is involved or infection diseases
; FILE REFERENCE: 2183-5136US
; CURRENT APPLICATION NUMBER: US/10/004,219B
; CURRENT FILING DATE: 2001-11-02
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1) (455)
; OTHER INFORMATION: /note="Human AMCase"
US-10-004-219B-14
Query Match 83.8%; Score 2089.5; DB 9; Length 455;
Best Local Similarity 81.3%; Pred. No. 9.5e-171;
Matches 369; Conservative 40; Mismatches 42; Indels 3; Gaps 1;
QY 1 YNLICVFTNWAQYRPLGSGFKPDDINPCLCTHLIYAFAGMONEITTIENNDVTLYKAFN 60
DB 1 YOLTCYFTNWAQYRPLGSGFKPDDINPCLCTHLIYAFAGMONEITTIENNDVTLYQAFN 60
QY 61 DLKRNKSLKTLTAIGGNWFGTAPFTTWVSTSONRQTFTTSVKELROYGFGDLDDWEY 120
DB 61 GLKNNKSLKTLTAIGGNWFGTAPFTTWVSTPENRQTFTTSVKELROYGFGDLDDWEY 120
QY 121 PGSRSGPPQDKHLFTVLVKEMREAFQEAIESNRPLMTAAVAGGISNIQAGYIPELS 180
DB 121 PGSRSGPPQDKHLFTVLVKEMREAFQEAIESNRPLMTAAVAGGISNIQAGYIPELS 180
QY 181 KYLDFHVTYDLHGSWEGYTGENSEPLYKYPTETGNSNAYLVNDVYNNYKNGGAPAEKLI 240
DB 181 QYLDYIHVTYDLHGSWEGYTGENSEPLYKYPTDTSNAYLVNDVYNNYKNGGAPAEKLI 240
QY 241 VGFPEYGHFTILRNPSDNGIGAPTSQDGPAGAYTRQAGFWAYYEICTFLRSGATEVWDAS 300
DB 241 VGFPTYGHNFILRNPSNTGIGAPTSQDGPAGAYTRQAGFWAYYEICTFLKNGATQWDAP 300
QY 301 QVVPYAYKANWLGNDYNIKSFSVKAQWLKQNNFGGAMIWAILDDFTGSCDQGFPLTS 360
DB 301 QVVPYAYOGNVWGYDNIKSFPDIKAQWLKHNKFGGAMIWAILDDFTGFCNQGFPLIS 360
QY 361 TLNKALGISTEGCTAPDVPSEPTTPPGSGSGSGSGSGSGGFCADKADGLYPVA 417
DB 361 TLNKALGLQASACTAPAPQPIETAPPSGSGSGSGSGSGSGGFCVAVRANGLYPVA 420
QY 418 DDNAFWQINGITYQHQHQAAGLVFDTSCNCCNP 451
DB 421 NNRNAPFWCHVNGVTYQQNCQAGLVFDTSCDCCNW 454

RESULT 4

US-10-004-219B-1

; Sequence 1, Application US/10004219B

; Publication No. US2003008741A1

; GENERAL INFORMATION:

; APPLICANT: Macrozyme

; APPLICANT: Aerts, Johannes M.F.G.

; APPLICANT: Boot, Rolf G.

; TITLE OF INVENTION: A mammalian mucinase, its recombinant production, and

; TITLE OF INVENTION: its use in therapy or prophylaxis against diseases in

; TITLE OF INVENTION: which mucus is involved or infection diseases

; FILE REFERENCE: 2183-5136US

; CURRENT APPLICATION NUMBER: US/10/004,219B

; CURRENT FILING DATE: 2001-11-02

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1

; LENGTH: 476

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: human AMCase

; OTHER INFORMATION: amino acid sequence deduced from cDNA sequence

US-10-004-219B-1

Query Match

Best Local Similarity 83.8%; Score 2089.5; DB 9; Length 476;

Matches 369; Conservative 40; Mismatches 42; Indels 3; Gaps 1;

QY 1

YNLCIVFTWAQYRPGLSGFKDDINPCLCTHLIYAFAGQNNNEITTIWNVDTLYKAFN 60

22 YQTCYFTWAQYRPGLSGFMEDNDPCCLCTHLIYAFAGRONNEITTIWNVDTLYQAFN 81

QY 61

DLKRNKSLKTLAIGGWNFGTAPFTTWSTQNRQTFITSVIKFLQYGFGLDLDWEY 120

Db 82

GLKRNKSLKTLAIGGWNFGTAPFTTWSTQNRQTFITSVIKFLQYGFGLDLDWEY 141

QY 121

PGSRGSPQDKHLFTVLVKEMREAFQEAIESNRPRMLMTAAVAGGISNIQAGYEIPELS 180

Db 142

PGSRGSPQDKHLFTVLVKEMREAFQEAIESNRPRMLMTAAVAGGISNIQAGYEIPELS 201

QY 181

KYLDFHVMYTDLHSGWEGYTGENSPLYKYPTETGNSNAYLNVDYVNNYKNGAPAEKLI 240

Db 202

QYLDYHVMYTDLHSGWEGYTGENSPLYKYPTETGNSNAYLNVDYVNNYKNGAPAEKLI 261

QY 241

VGPPEYGHFTILNPSDNGIGAPTSBGDPAGATROAGFWAYEICTFLRSGATEVWDAS 300

Db 262

VGPPTGHNFTILNPSDNGIGAPTSBGDPAGATROAGFWAYEICTFLRSGATEVWDAS 321

QY 301

QVVPYAYKANWGLYDNISFVSKAQLKQNNFGGAMIWALDIDFTGSCDQKFPPLTS 360

Db 322

QVVPYAYQVNVWYDNIKSFDIKAKQLKXKKEGGAMWALDIDFTGTCNQKPEPLIS 381

QY 361

TLNKALGISTEGCTAPDVPSEPVTPPP---GGSGGGSGSGSGSGSGGFCADKADGLYPVA 417

Db 382

TLKALGLASASCTAPQAPTEPIITAAPSGGNGSGSGSGSGSGGFCAVRANGLYPVA 441

QY 418

DORNAFMQINGITYOQHCQAGLVFTSCNCCNW 451

Db 442

NNRNFAMHCVNGVTYQONCOAGLVFTSCDCCNW 475

RESULT 5

US-10-004-219B-10

; Sequence 10, Application US/10004219B

; Publication No. US2003008741A1

; GENERAL INFORMATION:

; APPLICANT: Macrozyme

; APPLICANT: Aerts, Johannes M.F.G.

; APPLICANT: Boot, Rolf G.

; TITLE OF INVENTION: A mammalian mucinase, its recombinant production, and

; TITLE OF INVENTION: its use in therapy or prophylaxis against diseases in

; TITLE OF INVENTION: which mucus is involved or infection diseases

; FILE REFERENCE: 2183-5136US

; CURRENT APPLICATION NUMBER: US/10/004,219B

; CURRENT FILING DATE: 2001-11-02

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 10

; LENGTH: 445

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: SITE

; LOCATION: (1)..(445)

; OTHER INFORMATION: /note="Human chitotriosidase"

US-10-004-219B-10

Query Match

Best Local Similarity 52.2%; Score 1301.5; DB 9; Length 445;

Matches 235; Conservative 71; Mismatches 123; Indels 21; Gaps 3;

QY 3

LICYFTNWAQYRPGLSGFKDDINPCLCTHLIYAFAGQNNNEITTIWNVDTLYKAFNDL 62

Db 3

LVCYFTNWAQYRQGEARFLPKDLDPSLCTHLIYAFAGTNHQLSTTEWNDLTYQEFNGL 62

QY 63

KVRNSKLKTLAIGGWNFGTAPFTTWSTQNRQTFITSVIKFLQYGFGLDLDWEYFG 122

Db 63

KQONPKLTLAIGGWNFGTQKFTDMVATANNRQTFVNSAIRFLKYSFGGLDLDWEYFG 122

QY 123

SRGSPQDKHLFTVLVKEMREAFQEAIESNRPRMLMTAAVAGGISNIQAGYEIPELSKY 182

Db 123

SGSPAVDKRFTLLVQDLANAFQEAQTSKGERLLLSAAVPAGQTYVDAGYEVDKIAQN 182

QY 183

LDIFIHVMYTDLHSGWEGYTGENSPLYKYPTETGNSNAYLNVDYVNNYKNGAPAEKLIYG 242

Db 183

LDVNLMAVDHSGWEKVTGHSPLYKQESGAASLNVDAAVQWLQKGTPTASKLIIG 242

QY 243

PPEYGHFTILNPSDNGIGAPTSBGDPAGATROAGFWAYEICTFLRSGATEVWDASOE 302

Db 243

MPTYGRSTFLASSDTRVGAPATGSGTPGPTKEGMLAYEVCSSW--KGATKORIQQOK 300

QY 303

VPYAYKANWGLYDNISFVSKAQLKQNNFGGAMIWALDIDFTGSCDQKFPPLTSTL 362

Db 301

VPTTFRDQWQFDDVESFKTKVYLKQKGLGGAMWALDIDDFAGFCNQGRYPLIQT 360

QY 363

NKALGISTEGCTAPDV-----PSEPVTTPPGSGSGSGSGSGSGSGGFCADKADGLYP 415

Db 361

ROELSLPYLPSTPELVKPCQPSEP-----EHCPSPGQDTFCQKADGLYP 408

QY 416

VADDRNFMQINGITYOQHCQAGLVFTSCNCCNW 451

Db 409

NPRERSFVSCAAGRLFQOQSCPTGLVFSNSCKCCTW 444

RESULT 6

US-10-097-340-45

; Sequence 45, Application US/10097340

; Publication No. US20030087250A1

; GENERAL INFORMATION:

; APPLICANT: John MONAHAN

; APPLICANT: Manjula GANNAVAPARU

; APPLICANT: Sebastian HOERSCH

; APPLICANT: Shubhangi KAMATKAR

; APPLICANT: Steve G. KOVATS

; APPLICANT: Rachel E. MEYERS

; APPLICANT: Michael MORRISSEY

; APPLICANT: Peter OLANDT

; APPLICANT: Ami SEN

; APPLICANT: Peter VEIBY

; APPLICANT: Gordon B. MILLS

; APPLICANT: Robert C. BAST, Jr.

; APPLICANT: Karen LU

; APPLICANT: Rosemarie SCHMANDT

; APPLICANT: Xumei ZHAO

; APPLICANT: Karen GLATT


```

Qy 113 GLDLDWEYPGSR-GSPQDQKHLFTVLVKENRREAFQEAEIESNRPLRMVMTAAVAGGISNIQ 171
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 154 GLDLDWEYPGSLGNPKIDKQNTLALVRELKDAFEPHG-----YLLTAAVSPGDKID 206
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 172 AGYEIPELSKYLDPIHVMVTYDLHGSWEGYTGENSEPLYKYPTETGS-NAYLVNDVVMNWK 230
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 207 RAYDIIKELNKLFWMMVMVTYDYGHWENFYGHNAPLYKRPDEYDELHTYFNVNVTMHHYL 266,
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 231 NNGAPAEKLTIVGPPEYCHTFLRNPDSNGICAPTSGDGPAAGYTRQAGFWAYYEICTFLR 290
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 267 NNGATRDKLVMGVPPFYGRAWSIEDRSKJLGDPAKMSPPGFI-SGEBGVLSYIELCOLFQ 326
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 291 SGATEW----DASQEVPAKYKANEMLYGYNIIKSFSVKAQWLKQNNPFGGAMIWAIDLDDF 346
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 327 K--EEMHIQYDEYYNAPYGYNDKIWGYDDLASISCKLAFKLKELGSGVVMVMSLEND 383
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 347 TGSFCDOCKEPLTSTLKA-----LGISTEGCTAPDVPSEPVTP 386
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 384 KG-HCG-----PKNPLLNKVNMINGEKNSFECILGPSTTTPTTTPTTTPTT 433
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 10
US-10-218-743-18
; Sequence 18, Application US/10218743
; Publication No. US20030096779A1
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.
; APPLICANT: Hunter, Shirliev Wu

```

TITLE OF INVENTION: NOVEL DERIVATIVES OF NUCLEIC ACID NUCLEOTIDES, PROVIDING
 A METHOD OF PREPARING THE SAME, AND USES THEREOF
 FILE REFERENCE: AU-2-C3
 CURRENT APPLICATION NUMBER: US/10/218,743
 CURRENT FILING DATE: 2002-08-13
 PRIOR APPLICATION NUMBER: US/09/292,225
 PRIOR FILING DATE: 1999-04-15
 PRIOR APPLICATION NUMBER: 60/098,909
 PRIOR FILING DATE: 1998-09-02
 PRIOR APPLICATION NUMBER: 60/085,295
 PRIOR FILING DATE: 1998-05-13
 PRIOR APPLICATION NUMBER: 60/098,565
 PRIOR FILING DATE: 1998-04-17

```

; PRIOR APPLICATION NUMBER: 007096, 565
; PRIOR FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: 09/062, 013
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Dermatophagoides farinae
US-10-218-743-18

Query Match          29.3%;   Score 729.5;   DB 9;   Length 555;
Best Local Similarity 36.9%;   Pred No. 4.3e-54;
Matches 153; Conservative 71; Mismatches 144; Indels 47; Gaps 10;

Qy      3 LICYFTNWAQYRGLSGFKPDINPCLCTHLIYAFAGMQNNETTIEWN-----D 52
       :||| | : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db      35 IVCVVGWTSVYHK-VDPYTTIEDIPPKCTHLMYGFAKIDEVKYTVQFDDPNHNSWE 93

```

[illegible]


```

; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 371
; TYPE: PRT
; ORGANISM: Helianthus annuus
US-09-923-844B-2

Query Match 19.1%; Score 476; DB 9; Length 371;
Best Local Similarity 34.1%; Pred. No. 1.3e-32;
Matches 119; Conservative 53; Mismatches 135; Indels 42; Gaps 13;

QY 6 YFTNWAQ-YRPGSGSPKDDINPCLCTHLIYAFAGMNNNEIT---TIEWNDVTLYKAFND 61
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 30 YPWSWAQDFLP-----PSNIQTAYFTHVYAF--LSPNNVTQFDVHRTTASALNSFNT 81

QY 62 -LKNRSKLUKTLAIGWNFGTAP-FTTWSTVSQNRCTFTSIVKELQYGFQGLDLDWE 119
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 82 ALHGKPNPVKTFESIGGSAGVQKLFSLKSSPCGSRAAFIRSTIQVARNYYPFGADLDWE 141

QY 120 YPGSRGSPDKHLFTVLVKEMEAPEQEAIESNRRLMVTAAVA-----GGISNIOA 172
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 142 YPTQ-----TDMNFGLLDDEWAVVNEATSTGKPELLLSAATRIPEVRDNGV---A 193

QY 173 GYBIPELSKYLDFIHVMTYDLHGSW-EGYTGENSEPLYKYPTETGSNAYLNVDYVMNYWKN 231
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 194 KYPVASINKNLGDNAMCYDHYGFWPTPDATGAPAALY-----NPGSLSTSNGLQSWIS 247

QY 232 NGAPAEKLIUVFPEYGHFTFLRNPDSNGIGAPSGDGPGAGAYTRQAGFWAYYEICTF-LR 290
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 248 AGTQRQKLVMGMPLYGWTWKLKNPVSNGIGAPAAIGPG-----NEGAMLYSEVQQFNAQ 302

QY 291 SGATEVMDASQEVYAYKANEWLGYNIKSFSVKAOWLKONNFGGAMIW 339
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 303 NNARVVYDITQTSYYSYSGTTWIGYDDVNSVQRKVQAKSLNIGGFFW 351

RESULT 15
US-09-748-033-3
; Sequence 3, Application US/09748033
; Patent No. US20020069431A1
; GENERAL INFORMATION:
; APPLICANT: Broadway, Roxanne M.
; APPLICANT: Gongora, Carmenza E.
; TITLE OF INVENTION: EFFECT OF ENDOCHITINASE AND CHITOBIOCIDASE AND THEIR
; TITLE OF INVENTION: ENCODING GENES ON PLANT GROWTH AND DEVELOPMENT
; FILE REFERENCE: 19603/3091
; CURRENT APPLICATION NUMBER: US/09/748,033
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/172,003
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 376
; TYPE: PRT
; ORGANISM: Streptomyces albidoflavus
US-09-748-033-3

Query Match 10.6%; Score 265.5; DB 10; Length 376;
Best Local Similarity 25.9%; Pred. No. 1.4e-14;
Matches 108; Conservative 50; Mismatches 148; Indels 111; Gaps 24;

QY 6 YFTNWAQYRPL-----GSPKDDINPCLCTHLIYAFAGMNNNEITIE----- 49
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 14 YFTWGVYGRNVHVKNLVTSGSAEK-----ITHINYSFGNVQGGKCTIGDSFAAYDKA 66

QY 50 -----WNDVTLYKAFNDLKQRNSK---LKTLLAIGWNFGTAPFTTWYSTSQ 93
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 67 YTAESVDGVADTW-DOPLRGNFNQURKLKAKYPHIKVLSWFCGGTW-SGGFTDAV--K 121

QY 94 NROTFTITSIVKFLROYG-----FQGLDLDWEYPSGRG-----SPPODKHLFTVLVKEMRE 143

```

122	NPAAFAKSHDLVEDPRWADV	GDIGIDLDWEYPRNACGLSCDSSGPAALKH	----	WQVAMRA	177
144	AFEQBAIESNRPRMLWMTAA	VAGGISN----	IQAGYEIPELSKYLDFIHVMTYDLHGSWEGY	200	
178	QFGTD-----	LVTAAITADASSGKGLDAA-	DYAGAAQYFPDWNVTNYTFFFGAWD-K	226	
201	TCENSPLYKYPTETGSNAYLN	V----	VNMYWKQNGCAPAEKLI	VGFPYEGHTFILR	253
227	TGTPAP-----	HSALNSYSGIPKADFHSA	AAATAKLAKGVPA	SKLLGIGFVGRGW----	277
254	NPSDNGIGAPTSG----	DGPA-GATROAGFWAYYEI	CTFLRS	GATEWDA	SQBEVYAYKA 309
278	-----	TGVTQDPAGGTATGPA	TGTV--	EAGIEDYKVLKNTCPATG	VGGTA-----YAKCG 326
310	NEWLGYDNIKSFSVKAQWL	KQNNFGGAMIWALD	DDFTGSCFCDQGK	FPLTSTLNKAL	366
327	SNWNSYDTPATIKTKTKW	AKDQGLGGAFFW----	BFSG----	DTAGGELVSAMD	SGL 375

Search completed: June 29, 2003, 21:13:33
Job time : 18.3846 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 29, 2003, 21:03:33 ; Search time 2938 Seconds
(without alignments)
4477.355 Million cell updates/sec

Title: US-10-004-219b-9

Perfect score: 2493

Sequence: 1 YNLICYFTNWAQYRPLGSLF.....QQHCOAGLVFDTSCNCNWP 452

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -MODEL=frame+ p2n model -DEV=xl

-Q=/cgn2_1/USPTO_spool/US10004219/runat_24062003_160228_8159/app_query.fasta_1.1877
-DB=GenEmbl -QMT=fastap -SUFFIX=rg -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10004219 @CGN 1 1 6828 @runat_24062003_160228_8159 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMSOOT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.pln.*
36: em.htg.ram.*
37: em.htg.vrt.*
38: em.sv.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2493	100.0	1530	10 AF290003	AF290003 Mus muscu
2	2488	99.8	1526	10 AF154571	AF154571 Mus muscu
3	2488	99.8	1538	10 BC011134	BC011134 Mus muscu
4	2488	99.8	1557	10 BC034548	BC034548 Mus muscu
5	2089.5	83.8	1625	9 AF290004	AF290004 Homo sapi
6	1991.5	79.9	1529	4 AB051629	AB051629 Bos tauru
7	1778.5	71.3	1369	6 AX05989	AX05989 Sequence
8	1764.5	70.8	1354	9 AB025008	AB025008 Homo sapi
9	1446	58.0	1188	9 AB025009	AB025009 Homo sapi
10	1414.5	56.7	4250	9 AK098814	AK098814 Homo sapi
11	1397	55.6	1525	10 M94584	M94584 Mus musculu
12	1379	55.3	1506	10 D87757	D87757 Mus musculu
13	1373	55.1	1209	10 AY049765	AY049765 Mus muscu
14	1371	55.0	1562	5 BJA345054	BJA345054 Bufo japo
15	1368	54.9	1209	10 AY065557	AY065557 Mus muscu
16	1301.5	52.2	1633	9 HSU29615	U29615 Human chit
17	1301.5	52.2	1636	6 AR206041	AR206041 Sequence
18	1301.5	52.2	1636	6 AR212138	AR212138 Sequence
19	1301.5	52.2	1643	6 AX108750	AX108750 Sequence
20	1301.5	52.2	1643	6 AR172107	AR172107 Sequence
21	1295.5	52.0	1656	6 AR206042	AR206042 Sequence
22	1295.5	52.0	1656	6 AR212139	AR212139 Sequence
23	1295.5	52.0	1656	6 AX108752	AX108752 Sequence
24	1286	51.6	1710	9 HSU062862	U62862 Homo sapien
25	1286	51.6	1713	6 AR172108	AR172108 Sequence
26	1253.5	50.3	93740	9 AL513202	AL513202 Human DNA
27	1200	48.1	1599	9 AK055165	AK055165 Homo sapi
28	1069.5	42.9	233175	2 AC125461	AC125461 Mus muscu
29	1051	42.2	2504	6 E09046	E09046 cDNA encodi
30	1051	42.2	2525	10 MOSOGP	D32137 Mouse mRNa
31	1043	41.8	91569	9 AL356387	AL356387 Human DNA
32	1042.5	41.8	2353	10 NAU15048	U15048 Mesocricetu
33	1042.5	41.8	2387	10 HAMOGP	D32218 Hamster mRN
34	1033.5	41.5	2366	6 E09047	E09047 cDNA encodi
35	1014	40.7	6439	3 AB074977	AB074977 Haemaphys
36	1013	40.6	2198	9 HSU09550	U09550 Human ovidu
37	1010	40.5	1994	4 BOV95BOGP	D16639 Bovine mRNa
38	1010	40.5	1994	6 E09045	E09045 cDNA encodi
39	1009.5	40.5	2228	9 BABEDOSG	M59903 Papio hamad
40	1009.5	40.5	2237	9 MU087259	U87259 Macaca mula
41	1005.5	40.3	1406	9 AF484550	AF484550 Macaca ra
42	1005.5	40.3	1909	4 AF347052	AF347052 Oryctolag
43	1003	40.2	2034	4 OAU16719	U16719 Ovis aries
44	1001	40.2	1596	4 OAU17988	U17988 Ovis aries
45	995	39.9	3946	3 AY051988	AY051988 Drosophil

ALIGNMENTS

RESULT 1

AF290003
LOCUS AF290003 1530 bp mRNA linear ROD 27-FEB-2001
DEFINITION Mus musculus acidic mammalian chitinase precursor, mRNA, complete cds.
ACCESSION AF290003.1 GI:12597290
VERSION AF290003.1
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1530)
AUTHORS Boot, R.G., Bloomaart, E.F., Swart, E., Chauharali-van der Viugt, K., Bijl, N., Moe, C., Place, A., and Aerts, J.M.
TITLE Identification of a novel acidic mammalian chitinase distinct from chitotriosidase
J. Biol. Chem. 276 (9), 6770-6778 (2001)
MEDLINE 21125893
PUBMED 11085997
REFERENCE 2 (bases 1 to 1530)
AUTHORS Boot, R.G., Verhoeck, M., Swart, E. and Aerts, J.M.
TITLE Direct Submission
JOURNAL Submitted (26-JUL-2000) Dept. of Biochemistry, Academic Medical Center, University of Amsterdam, Meibergdreef 15, Amsterdam 1105 AZ, The Netherlands
FEATURES
source
1..1530
/organism="Mus musculus"
/strain="BALB/c"
/db_xref="taxon:10090"
/sex="male"
/tissue type="lung"
/dev stage="9-11 weeks"
3..1424
/BC_number="3.2.1.14"
/note="chitinase precursor; AMCase"
/codon start=1
/product="acidic mammalian chitinase precursor"
/protein_id="AAG60018.1"
/db_xref="GI:12597291"
translation="MAKLLVLTGLALLNAQLGSAYNLICYFTNWAQVRPGLGSKPKD
DINPLCTHLIYAFAGMKNETITIEWNDVTLYKAFNDLNKNSKLTLALAGWNFG
TAPFTVSTQNRQFTITSVKFLQYDGLDLDMEYPGSRGSPQDRLHFLTLVK
ENREAFQEARISNRPRLMTAAVAGGINSIOAGVEIPSLKYLDF IHVMTVDLHGSW
EGVTGNSPLVYPTETGNSYLVNDVYNNYKNGCAPAEKLI VGFPEVGHFTILRNP
SDNGIGAPTSQGPAGAVTQAGFWAYEICTFLSGATEVWDASQEPVYAKNEWL
GTDNKSFSVKQWLKQNNFGGAMWAILDFTGSCDQGFPLTSLINKALGISTE
GCTAPDVPSEPTVTPPGSGSGSGSGSGSGSGGFCADKADGLYPVADDRNPFQWQING
ITYQHCQAGLVFDTSNCNCP"

BASE COUNT 382 a 383 c 366 g 399 t
ORIGIN

Alignment Scores:
Pred. No.: 2.5e-178 Length: 1530
Score: 2493.00 Matches: 452
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-10-004-219B-9 (1-452) x AF290003 (1-1530)

QY 1 TyrAsnLeuileCysTyrPheThrAsnTpaAlaGlnTyrArgProGlyLeuGlySerPhe 20
Db 66 TACAATCTGATATGCTATTTCACCAACTGGGCCAGTATCGGCCAGGTCTGGGAGCTTC 125
QY 21 LysProAspAspIleAsnProCysLeuCysThrHisLeuileTyrAlaPheAlaGlyMet 40
Db 126 AAGCGTATGATGACATAACCCCTGCTGTACTCACCTGATCTATGCTTCTGCGGATG 185
QY 4i GinAsnAsnGluileThrThrileGluThrAsnAspValThrleuTyrlysalapheAsn 60
Db 186 CAGAACATGATGATCACCACCATAGAAATGGAATGATGTTACTCTCTATAAAGCTTTCAAT 245

QY 61 AspLeuLysAsnArgAsnSerLysLeuLysThrLeuLeuAlaileGlyGlyTyrAsnPhe 80
Db 246 GACTTGAATAAAG 305
QY 81 GlyThrAlaProPheThrMetValSerThrSerGlnAsnArgGlnThrPheleThr 100
Db 306 GGAACTGCTCTCTTCACTACCATGCTTTCCACTTCTCAGAACCCGACAGACCTTCATTACC 365
QY 101 SerValileLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyr 120
Db 366 TCAGTCATCAAAATTTTCGCTCAGTATGGTGTGATGAGTGGAGCTGGAGCTGGAGTAC 425
QY 121 ProGlySerArgGlySerProGlnAspLysHisLeuPheThrValLeuVallyGlu 140
Db 426 CCAGGCTCAGTGGGAGGCTCTCTCAGGACAGCATCTCTTCACCTGTCTGTGTGAGGNA 485
QY 141 MetArgGluAlaPheGluGlnGluAlaileGluSerAsnArgProArgLeuMetValThr 160
Db 486 ATGCGTGAAGCTTTTTCAGCAGGAGGCTATTGAGAGCAACAGGCCAGACTGATGTTACT 545
QY 161 AlaAlaValAlaGlyLysSerAsnileGlnAlaGlyTyrGluileProGluLeuSer 180
Db 546 GCTGCTGTAGCTGGTGGGATTTCCAAACATCCAGGCTGGCTATGAGATCCCTGAACCTTCT 605
QY 181 LysTyrLeuAspPheileHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr 200
Db 606 AGTACCTGGATTTCATGTCATGTCATGATGATGATGATGATGATGATGATGATGATGAT 665
QY 201 ThrGlyGluAsnSerProLeuTyrlyThrProThrGluThrGlySerAsnAlaTyrLeu 220
Db 666 ACTGGGAGAGATAGTCTCTTTACAAATACCTACTAGAGCTGGTAGCAATGCCCTACCTC 725
QY 221 AsnValAspTyrValMetAsnTyrTrpLysAsnAsnGlyAlaProAlaGluLysLeuile 240
Db 726 AATGGGATATTGTCATGAATATTGGAGAGCAATATGAGGCCAGCTGAGAGAGCTCAAT 785
QY 241 ValGlyPheProGluTyrGlyHisThrPheileLeuArgAsnProSerAspAsnGlyIle 260
Db 786 GTTGGATTTCCAGAGATAGGACACACCTTCATCTCAGAGAACCCCTCTGATANTGGAAT 845
QY 261 GlyAlaProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrp 280
Db 846 GGTGCCCTTACCTCTGTGTGATGGCTCTGTGGCTCTGTGGCTCTGTGGCTCTGTGGCT 905
QY 281 AlaTyrTyrGluileCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSer 300
Db 906 GCCTACTATGAGATTTCACCTTTCTGAGAAAGTGGAGCCACTGAGGTCTGGGATGCCTCC 965
QY 301 GlnGluValProTyrAlaTyrLysAlaAsnGluTrpLeuGlyTyrAspAsnileLysSer 320
Db 966 CAAGAGTGGCTATGCTTAAAGCCCAAGCAGTGGCTTGGCTATGACAAATATCAAGAGC 1025
QY 321 PheSerValLysAlaGlnTrpLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTrpAla 340
Db 1026 TTCAGTGTAAAGGCTCAGTGGCTTAAAGCAGAACAAATTTGGAGGTGCCATGATCTGGGCC 1085
QY 341 IleAspLeuAspAspPheThrGlySerPheCysAspGlnGlyLysPheProleuThrSer 360
Db 1086 ATTGACCTTGATGACTTCACTGGCTCTTCTGTGATCAGGGGAAATTTCCCTGTGACTTCT 1145
QY 361 ThrLeuAsnLysAlaLeuGlyLysSerThrGluGlyCysThrAlaProAspValProSer 380
Db 1146 ACTTTGAACAAAGCCCTTGGCATATCCACTGAGGTTGCCAGCTCTCTGAGCTGCTTCC 1205
QY 381 GluProValThrThrProGlySerGlySerGlyGlyGlySerSerGlyGlySerSer 400
Db 1206 GAGCCAGTACTTCTCTCCAGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTCT 1265
QY 401 GlyGlySerGlyPheCysAlaAspLysAlaAspGlyLeuTyrProValAlaAspAspArg 420
Db 1266 GGAGCAGTGGATTCTGTGGCAGAACAGATGAGCTTACCTTGTGGCAGATGACAGA 1325


```
QY 341 IleAspLeuAspAspPheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSer 360
DB 1081 ATTGACCTTGATGACTTCACTGGCTCTTCTGTGATCAGGAAATTTCTCTGACTTCT 1140
QY 361 ThrLeuAsnLysAlaLeuGlyIleSerThrGluGlyCysThrAlaProAspValProSer 380
DB 1141 ACTTTGAACAAGCCCTTGGCATATCCACTGAAGGTTGCACAGCTCTCTGACGTCCTCC 1200
QY 381 GluProValThrThrProProGlySerGlySerGlyGlySerGlySerGlySer 400
DB 1201 GAGCCAGTGACTACTCTCTCAGGAAGTGGAGTGGGGTGGAGCTCCGGAGAGCTCT 1260
QY 401 GlyGlySerGlyPheCysAlaAspLysAlaAspGlyLeuTy-ProValAlaAspAspArg 420
DB 1261 GGAGGAGTGGATCTGTGGCAGCAAGCAGATGGCTCTACCTCTGGCAGATGACAGA 1320
QY 421 AsnAlaPheTrpGlnCysIleAsnGlyIleThrTyThrGlnGlnHisCysGlnAlaGlyLeu 440
DB 1321 AATGCTTTTGGCAGTGCATCAATGGAATCACATACAGCAGCATTTGTCAAGCAGGCTT 1380
QY 441 ValPheAspThrSerCysAsnCysCysAsnTrpPro 452
DB 1381 GTTTTGTATACAGCTGTAATTGCTGCAATGGCCA 1416

RESULT 3
BC011134
LOCUS BC011134 1538 bp mRNA linear ROD 07-AUG-2002
DEFINITION Mus musculus, Similar to eosinophil chemotactic cytokine, clone
ACCESSION MGC:18771 IMAGE:4165150, mRNA, complete cds.
VERSION BC011134
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (25-JUL-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgabbs@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: ang@bcm.tmc.edu
Gunnaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK plate: 24 Row: e Column: 11.
FEATURES
Location/Qualifiers
1..1538
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="FVB/N"
/clone="MGC:18771 IMAGE:4165150"
/tissue type="Sallivary gland, 10 week old female mouse"
/clone lib="NCI CGAP_SG2"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
323..1420
/codon_start=1
```

```
/product="Similar to eosinophil chemotactic cytokine"
/protein id="AAH11134.1"
/db_xref="GI:15029822"
/translation="MVSISQNRQTFITSVIKELRQYGFDFGLDLDWEYSGSRGSPQDDK
HLFTVLVKEMREAFQEAEIESNRPLMVTAAVAGISNIQAGYEIPELUSKYLDLFIHVM
TYDLHGSVEGTGNSPLTKPTETGSNAYLNVDYVMNKNNGNAPAEKLIYGFPEYG
HTFILRNPDSNGIGATSGDGPAGPYTRQAGFWAYEICTFLRSAGTAVWDAQSPVY
AKVNEWLGYNIKISFSKQAWLKNNGFGGAMIWAILDLDFTGSCDQKPLTSLN
KALGISTEGCTAPDPVSEPTVTPGSSGSGSSGSGSGSCGFCADKADGLYFVADRN
AFWOCINGITQQHCQAGLVDFDTCNCNWP"
BASE COUNT 393 a 382 c 365 g 398 t
ORIGIN
Alignment Scores:
Pred. No.: 5 97e-178 Length: 1538
Score: 2488.00 Matches: 451
Percent Similarity: 99.78% Conservative: 0
Best Local Similarity: 99.78% Mismatches: 1
Query Match: 99.80% Indels: 0
DB: 10 Gaps: 0
US-10-004-219B-9 (1-452) x BC011134 (1-1538)
QY 1 TyrAsnLeuIleCysTyThrPheThrAsnTrpAlaGlnTyArgProGlyLeuGlySerPhe 20
DB 62 TACAATCTGATATGCTATTTTCAACCACTGGGCCAGTATCGCCAGGTCCTGGGAGCTTC 121
QY 21 LysProAspAspIleAsnProCysLeuCysThrHisLeuIleTyThrAlaPheAlaGlyMet 40
DB 122 AAGCCTGATGACATTAACCCCTGCTGTACTACCTGATCTATGCTGCTGCTGCTGATG 181
QY 41 GlnAsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyLysAlaPheAsn 60
DB 182 CAGAACATGATGATCACCACCATAGATGATGATGATGATGATGATGATGATGATGATGAT 241
QY 61 AspLeuLysAsnArgAsnSerLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPhe 80
DB 242 GACTTGAAAAACAGAACAGCAAACTGAAACCCCTCTCTGGCAATTTGGAGGCTGGAACCTT 301
QY 81 GlyThrAlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheIleThr 100
DB 302 GGAACCTGCTCTCTTCACTACCATGCTTTTCCACTCTCTCAGAACCCGACAGCTTCATTACC 361
QY 101 SerValIleLysPheLeuArgGlnTyThrGlyPheAspGlyLeuAspLeuAspTrpGluTy 120
DB 362 TCAGTCATCAAAATTTCTGCGTCAGTATGGGTTGATGAGCTGGACCTGGACTGGGAATAC 421
QY 121 ProGlySerArgGlySerProGlnAspLysHisLeuPheThrValLeuValLysGlu 140
DB 422 CCAGGCTCAGCTGGGAGCCCTCTCAGGACCAAGCATCTCTTCACTGCTCTGCTGGTGAAGAA 481
QY 141 MetArgGluAlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThr 160
DB 482 ATGCGTGAAGCTTTTGGAGCAGGAGCTATTGAGAGCAACAGGCCCCAGACTGATGTTACT 541
QY 161 AlaAlaValAlaGlyGlyIleSerAsnIleGlnAlaGlyTyThrGluIleProGluLeuSer 180
DB 542 GCTGCTGATGCTGGTGGGATTTCCACATCCAGGCTGGCTGATGAGATCCCTGAACTTCT 601
QY 181 LysTyThrLeuAspPheIleHisValMetThrTyThrAspLeuHisGlySerTrpGluGlyTy 200
DB 602 AAGTACCTGGATTTTCATCCATGTCATGATACCTATGCTCCATGCTCTCTGGGAGGCTAC 661
QY 201 ThrGlyGluAsnSerProLeuTyThrTyThrProThrGluThrGlySerAsnAlaTyThr 220
DB 662 ACTGGGAGAAATAGTCTCTTTTACAAATACCTACTAGAGCTGGTAGCAATGCTTACCTC 721
QY 221 AsnValAspTyThrValMetAsnTyThrLysAsnAsnGlyAlaProAlaGluLysLeuIle 240
DB 722 AATGTGGATTATGTCATGAACTATTGGAGAACCAATGGAGCCCCAGCTGAGAGCTCATT 781
QY 241 ValGlyPheProGluTyThrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIle 260
```

```
Db 782 GTTGGATCCAGAGTATGACACACACCTTCATCTCTGAGAAACCCCTCTGATATATGGAATT 841
Qy 261 GlyAlaProThrSerGlyAspGlyProAlaGlyAlaThrArgGlnAlaGlyPheTrp 280
Db 842 GTGCCCCCTACTCTGGTGATGGCCCTGCTGGCCCTATACACAGACAGCTGGTTCGG 901
Qy 281 AlaTyrTyrGluLeuCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSer 300
Db 902 GCCTACTATGAGATTGACACCTTCTGAGAACTGGAGCCACTGAGCTCGGATGCTCC 961
Qy 301 GlnGluValProTyrAlaTyrIlysAlaGlnTyrLeuGlyTyrAspAsnIleYsSer 320
Db 962 CAAGAAGTGCCTATAGCCCTATAGGCCAACAGAGTGGCTTGGCTATGACAAATATCAAGAGC 1021
Qy 321 PheSerValIysAlaGlnTrpLeuIysGlnAsnAspPheGlyGlyValaMetIleTrpAla 340
Db 1022 TTCAGTGTAAAGCTCAGTGGCTTAAGCAGACAAATTTGGAGGTGCCATGATCGGGCC 1081
Qy 341 IleAspLeuAspAspPheThrGlySerPheCysAspGlnGlyIysPheProLeuThrSer 360
Db 1082 ATTGACCTTGATGACTTCACTGGCTCTTCTGTGATCAGGAAATTTCTCTGACTTCT 1141
Qy 361 ThrLeuAsnIysAlaLeuGlyIleSerThrGluGlyCysThrAlaProAspValProSer 380
Db 1142 ACTTTGAACAAAGCCCTTGGCATATCCACTGAAGGTTGCACAGCTCTCGACGTCTCC 1201
Qy 381 GluProValThrProProGlySerGlySerGlyGlyGlySerSerGlyGlySerSer 400
Db 1202 GAGCCAGTACTACTCTCCAGAAAGTGGAGTGGGGTGGAGCTCCCGAGAGAGCTCT 1261
Qy 401 GlyGlySerGlyPheCysAlaAspIysAlaAspGlyLeuTyrProValAlaAspAspArg 420
Db 1262 GGAGGAGTGGATTCTGTGCCGACAAAGCAGATGGCTCTACCTGTGGCAGATGACAGA 1321
Qy 421 AsnAlaPheTrpGlnCysIleAsnGlyIleThrTyrGlnGlnHisCysGlnAlaGlyLeu 440
Db 1322 AATGCTTTTGGCAGTGCATCAATGAATACATACACAGCAGCATTTGTCAAGCGGCTT 1381
Qy 441 ValPheAspThrSerCysAsnCysCysAsnTrpPro 452
Db 1382 GTTTTGTATACAGCTGTAAATTTGCTGCAACTGGCCA 1417
RESULT 4
LOCUS BC034548 1557 bp mRNA linear ROD 07-AUG-2002
DEFINITION Mus musculus, chitinase, acidic, clone MGC:19045 IMAGE:4189080,
mRNA, complete cds.
ACCESSION BC034548
VERSION BC034548.1 GI:21961190
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1557)
Strausberg, R.
Direct Submission
Submitted (24-JUL-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunnaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
```

```
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAP Plate: 24 Row: h Column: 16
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not
identity to protein.
FEATURES
Location/Qualifiers
source 1..1557
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="FVB/N"
/clone="MGC:19045 IMAGE:4189080"
/tissue type="Salivary gland, 10 week old female mouse"
/clone_lib="NCI CGAP_SG2"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
4..1425
/codon_start=1
/product="chitinase, acidic"
/protein_id="AAH34548.1"
/db_xref="GI:21961191"
/db_xref="LocusID:81600"
/translation="MAKLLLVTLALLLNAQLGSAYNLICFTYNTWAQYRPGLSFKPD
DINPLCLTHLYAPAGMONEITTEWNDVLYKAFNDLKNRNSKLTLLAIGGNVFG
TAPTWYTSQNRQETTSVKELROYGFDGLDLDWEVPGSRGSPDDKHLPTVLVK
EMREAFQEAIESNRPLMVAANVAGGSIHQAGYEIPELSKYLDHFHWTYDLNGSW
EGYTGNSPLKYIPETGNSVNLVNVYNNYKNGGAPAEKLIYGFPEYGHFTILRNP
SDNGIGAPTSKDGPGPYTRQAGFWAYEICTFLRSAGTEVMDASQEPYAYKANEWL
GYDNIKSPVKAQWLKQNNFGGAMIWALDDFTGSGCDQKFPILTSTLNKALGISTE
GCTAPDVPSEPTTPPGSGSGSGSGSGSGSGGFCADKADGLYIPVADDRNAPMQCING
ITYQHCQAGLVFDTSCNCWP"
BASE COUNT 410 a 382 c 366 g 399 t
ORIGIN
Alignment Scores:
Pred. No.: 6,06e-178 Length: 1557
Score: 2488.00 Matches: 451
Percent Similarity: 99.78% Conservative: 0
Best Local Similarity: 99.78% Mismatches: 1
Query Match: 99.80% Indels: 0
DB: 10 Gaps: 0
US-10-004-219B-9 (1-452) x BC034548 (1-1557)
Qy 1 TyrAsnLeuIleCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPhe 20
Db 67 TACAATCTGATATGCTATTTCACCAACTGGGCCAGTATCGGCCAGGTCTGGGAGCTTC 126
Qy 21 LysProAspAspIleAsnProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet 40
Db 127 AAGCCTGATGACATTAAACCCCTCGCTGTGTACTACCTGATCTATGCTCTTCTCGGATG 186
Qy 41 GlnAsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrIysAlaPheAsn 60
Db 187 CAGACATAGATGATCACCACCATAGATGATGATGATGATGATGATGATGATGATGATG 246
Qy 61 AspLeuIysAsnArgAsnSerIysLeuIysThrLeuLeuAlaIleGlyClyTrpAsnPhe 80
Db 247 GACTTGAAGAAACAGACAGCAACAAACCTCTCGGCAATTTGGAGCTTGGAACTTT 306
Qy 81 GlyThrAlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheIleThr 100
Db 307 GGAAGTCTCTCTTTCACATACCATGCTTTCCATCTTCAGAACCCGACACCTTCATTACC 366
Qy 101 SerValIleIysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyr 120
Db 367 TCAGTCATCAAAATTTCTGGTCAGTATGGTTTGTATGACTGGACTGGAGATAC 426
121 ProGlySerArgGlySerProProGlnAspIysHisLeuPheThrValLeuValIysGlu 140
```

```
Db 427 CCAGGCTCAGTGGGAGCCCTCTCAGGACAGCATCTCTCTCAGCTGTCTGGTGAAGGAA 486
Qy 141 MetArgGluAlaPheGluGlnGluAlaLeuSerAspArgProArgLeuMetValThr 160
Db 487 ATGCGTGAAGCTTTTGAGCAGGAGGCTATTGAGACACAGGCCAGACTGATGGTTACT 546
Qy 161 AlaAlaValAlaGlyGlyIleSerAsnIleGlnAlaGlyTyrGluIleProGluLeuSer 180
Db 547 GCTGCTGTAGTGGTGGATTTCCACATCCAGGCTGGCTATGAGATCCTGAACCTTCT 606
Qy 181 LysTyrLeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr 200
Db 607 AAGTACCTGGATTTCATCCATGTCATGACATATGACCTCATGGCTCCCTGGGAGGGCTAC 666
Qy 201 ThrGlyGluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeu 220
Db 667 ACTGGGGAGANATGCTCTTTTACAATATCCCTACTGAGACTGGTAGCAATGCTTACCTC 726
Qy 221 AsnValAspTyrValMetAsnTyrTrpLysAsnAsnGlyAlaProAlaGluLysLeuIle 240
Db 727 AATGTGGATTATGTATGTAACATTATGGAAGAACAAATGGAGCCCGAGCTGAGAAGCTCAT 786
Qy 241 ValGlyPheProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIle 260
Db 787 GTTGGATTCCAGAGTATGACACACACCTTCATCTGAGAAACCCCTCTGATATGGAATT 846
Qy 261 GlyAlaProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrp 280
Db 847 GGTGCCCTTACCTCTGTGTATGGCTCTGTGGCCCTATACACAGAGCTGGTGTCTGG 906
Qy 281 AlaTyrTyrGluIleCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSer 300
Db 907 GCTACTATGAGATTGACCTTCTTGAGAGTGGAGCCACTGAGGTCTGGGATGCTCTCC 966
Qy 301 GlnGluValProTyrAlaTyrLysAlaAsnGluTrpLeuGlyTyrAspAsnIleLysSer 320
Db 967 CAAGAAGTCCCTATGCTTATAGGCCAAAGCAAGTGGCTTGGCTATGACAAATATCAAGAGC 1026
Qy 321 PheSerValIleAlaGlnTrpLeuLysGlnAsnAsnPheGlyAlaMetIleTrpAla 340
Db 1027 TTCAGTGTAAAGCTCAGTGGCTTAAAGCAGAACAAATTTGGAGGTGCCATGATCTGGGCC 1086
Qy 341 IleAspLeuAspAspPheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSer 360
Db 1087 ATTGACCTTGTATGACTTCACTGGCTCTTCTGTGATCAGGGAATTTCTCTGACTTCT 1146
Qy 361 ThrLeuAsnLysAlaLeuGlyIleSerThrGluGlyCysThrAlaProAspValProSer 380
Db 1147 ACTTTGAACAAAGCCCTTGGCATATCCACTGAAGGTTGCACAGCTCCTGACGTGGCTTCC 1206
Qy 381 GluProValThrProGlySerGlySerGlyGlySerSerGlyGlySerSer 400
Db 1207 GAGCCAGTACTACTCTCTCAGGAGTGGAGTGGGGGTGGAGCTCCGAGGAGCTCT 1266
Qy 401 GlyGlySerGlyPheCysAlaAspLysAlaAspGlyLeuTyrProValAlaAspAspArg 420
Db 1267 CGAGGCAGTGGATTCTGTGCCGACAAAGCAGATGGCTCTTACCTGTGGCAGATGACAGA 1326
Qy 421 AsnAlaPheTrpGlnCysIleAsnGlyIleThrTyrGlnGlnHisCysGlnAlaGlyLeu 440
Db 1327 AATGCTTTTGGCAGTGCATCAATGGAATCACATACACAGCAGCATTTGTCAAGCAGGCTT 1386
Qy 441 ValPheAspThrSerCysAsnCysCysAsnTrpPro 452
Db 1387 GTTTTGTATACAGCTGTAAATGCTGCACACTGGCCA 1422
RESULT 5
AF290004
LOCUS
DEFINITION Homo sapiens acidic mammalian chitinase precursor, mRNA, complete
cDS
ACCESSION AF290004
VERSION AF290004.1 GI:12597292
```

```
KEYWORDS Homo sapiens.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1625)
AUTHORS Boot,R.G., Blommaert,E.F., Swart,E., Ghauharali-van der Vlugt,K.,
Bijsl,N., Moe,C., Place,A. and Aerts,J.M.
TITLE Identification of a novel acidic mammalian chitinase distinct from
chitinotrioidase
JOURNAL J. Biol. Chem. 276 (9), 6770-6778 (2001)
MEDLINE 21125893
PUBMED 11085997
REFERENCE 2 (bases 1 to 1625)
AUTHORS Boot,R.G., Verhoek,M., Swart,E. and Aerts,J.M.
TITLE Direct Submission
JOURNAL Submitted (26-JUN-2000) Dept. of Biochemistry, Academic Medical
Center, University of Amsterdam, Meibergdreef 15, Amsterdam 1105
AZ, The Netherlands
FEATURES
source
1..1625
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="stomach; lung"
104..1534
/EC_number="3.2.1.14"
/note="chitinase precursor; AMCase; chitinase activity
demonstrated by expression in COS cells; similar to
sequences presented in GenBank Accession Numbers AB025008
and AB025009"
/codon_start=1
/product="acidic mammalian chitinase precursor"
/protein_id="AAG60019.1"
/db_xref="GI:12597293"
/translation="MTKLILTLGLVLIINLQLSAYQLTCTFTWAQYRPGLRMPD
NIDPLCTLLIYAFAGRONNEITTEWDLTYQAFNGLKNKNSOLKTLAIGWNFG
TAPFTAMYSTPENQFTITSVKFLROYEPLDGFDEWYPGSRGSPQDKHLFTLVQ
EMRFAEQEAKQINKPRLMTAAVAAGISNIQSGVEIPLSQYLDYIHMTYDLHGSW
EGYTGENSLVKYPTDTSNAYLVNVDYVMYKONGAPAEKLIYGFPTYGHNFLSNP
SNTGKISDTSAGAGPAGPKESGIWAYVEICTFLKNGATCGMDAPQEPVYAYQGNVW
GYNIGKSFDIKAOWLKHNFEGAMVWALDLDFTGTCTFCNGKFPFLITLKALQLQSA
SCTPAQPIETPTAAPSNGSGSSSGSGSGGSCFCAVRANGLYPVANRRNFAFWHC
VNGVTYQNCQAGLVFDFTSCDCNWA"
sig_peptide 387 a 468 c 399 g 371 t
BASE COUNT 387 a 468 c 399 g 371 t
ORIGIN
Alignment Scores:
Score: 5,46e-148 Length: 1625
Pred. No.: 2089,50 Matches: 369
Percent Similarity: 90.0% Conservative: 40
Best Local Similarity: 81.2% Mismatches: 42
Query Match: 83.8% Indels: 3
DB: 9 Gaps: 1
US-10-004-219B-9 (1-452) x AF290004 (1-1625)
Qy 1 TyrAsnLeuIleCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPhe 20
Db 167 TACCAGCTGACATGCTACTTCCACCAACTGGGCCAGTACCGGCCAGGCTGGGGCGCTTC 226
Qy 21 LysProAspAspIleAsnProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet 40
Db 227 ATGCTTGACACATCGACCCCTGCTCTGTACCCACCTGATCTACGCCTTGTGGGAGG 286
Qy 41 GlnAsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrLysAlaPheAsn 60
Db 287 CAGAACACAGAGATCACCACTCGAATGGAACGATGTGACTCTCTACCACTTTCAT 346
Qy 61 AspLeuLysAsnArgAsnSerLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPhe 80
Db 347 GGCCTGAAAAATAGAACAGCCAGCTGANAACCTCTCTGGCCATTTGGAGGCTGGAACTTC 406
```


QY 81 GlyThrAlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheIleThr 100
 DB 407 GGGACTGGCCCTTTCACTGCCATGTTTCTACTCTGAGAACCGCAGACTTTTCATCAC 466
 QY 101 SerValIleLysPheLeuArgGlnThrGlyPheAspGlyLeuAspLeuAspTrpGluThr 120
 DB 467 TCAGTCATCAAAATCTCGGCCAGTATGAGTTTGAAGGCTGGACTTTGACTGGGAGTAC 526
 QY 121 ProGlySerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValLysGlu 140
 DB 527 CCTGGCTCTGGGGAGCCCTCTCAGACAGCATCTTCTACTGTCTGTGGCAGGAA 586
 QY 141 MetArgGluAlaPheGluGlnGluAlaLeuSerAsnArgProArgLeuMetValThr 160
 DB 587 ATGCGTGAAGCTTTTGGCAGGAGCCAGCAGCATCAACAGCCCGAGGCTGATGTCCT 646
 QY 161 AlaAlaValAlaGlyGlyLeuSerAsnIleGlnAlaGlyTyrGluLeuProGluLeuSer 180
 DB 647 GCTGAGTAGTCTGGCATCTCCAAATATCCAGTCTGGCTATGAGATCCCCCACTGTCA 706
 QY 181 LysTyrLeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr 200
 DB 707 CAGTACCTGGACTATCATGATCATGACCTAGACCTCATGGCTCTCTGGAGGGCTAC 766
 QY 201 ThrGlyGluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeu 220
 DB 767 ACTGGAGAGAACAGCCCTCTTACAAATACCCGAGTCTGACACCGCAGCAGCCCTACCTC 826
 QY 221 AsnValAspTyrValMetAsnTyrTrpLysAsnAsnGlyAlaProAlaGluLysLeu 240
 DB 827 AATGTGGATTATGTATGAATCTACTGGAAGGCAATGGAGCACCACTGAGAAGCTCATC 886
 QY 241 ValGlyPheProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIle 260
 DB 887 GTTGATTCCTTACCTATGAGCAACACTTCTCCTGAGCAACCCCTCCACACTGGAAT 946
 QY 261 GlyAlaProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrp 280
 DB 947 GGTGCCCCCACCCTCTGGTCTGCTGCTGGCCCTATGCCAAGGAGTCTGGGATCTGG 1006
 QY 281 AlaTyrTyrGluIleCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSer 300
 DB 1007 GCTTACTAGAGATCTGTACTCTCTGAAATAATGGAGCCACTCAGGGATGGATGCCCT 1066
 QY 301 GlnGluValProTyrAlaTyrLysAlaGlnTyrLeuGlyTyrAspAsnIleLysSer 320
 DB 1067 CAGGAAGTCCCTTATGCTTATCAGGCAATGTGGTGGTGGTGTATGACAACTCAAGAGC 1126
 QY 321 PheSerValLysAlaGlnTrpLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTrpAla 340
 DB 1127 TTCGATATTAAAGCTCAATGGCTTAAGCACACAAATTTGGAGGCCCATGCTCTGGGC 1186
 QY 341 IleAspLeuAspAspPheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSer 360
 DB 1187 ATTGATCTGGATGACTTCACTGGCACTTCTGCAACACAGGCAAGTTTCCCTTAATCTCC 1246
 QY 361 ThrLeuAsnLysAlaLeuGlyIleSerThrGluGlyCysThrAlaProAspValProSer 380
 DB 1247 ACCCTGAAGAGGCCCTCGGCTCGAGTGAAGTTCACGGCTCCAGCTCAGGCCATTT 1306
 QY 381 GluProValThrThrProProProProProProProProProProProProProProPro 397
 DB 1307 GAGCCATAAATCTGCTCTCCAGTGGCAGCGGAGCGGAGCGGAGTACGAGCTCTGA 1366
 QY 398 GlySerSerGlyGlySerGlyPheCysAlaAspLysAlaAspGlyLeuTyrProValAla 417
 DB 1367 GGCAGCTCGGAGGAGTGGATCTGTGCTGTGAGAGCCCAAGCGGCTCTACCCCGTGA 1426
 QY 418 AspAspArgAsnAlaPheThrGlnCysIleAsnGlyIleThrTyrGlnGlnHisCysGln 437
 DB 1427 AATAACAGAAATGCTCTCTGGCAGTCTGGGATGGAGTACGCTACAGCAGAGTCCAG 1486
 QY 438 AlaGlyLeuValPheAspThrSerCysAsnGlyCysAsnTrp 451

DB 1487 GCCGGCTGTCTTCGACACACGAGCTGTGATTGCTGCAACTGG 1528
 RESULT 6
 LOCUS AB051629 1529 bp mRNA linear MAM 07-NOV-2001
 DEFINITION Bos taurus cbp b04 mRNA for chitin binding protein b04, complete cds.
 ACCESSION AB051629 GI:16754894
 VERSION AB051629.1
 KEYWORDS Bos taurus liver cDNA to mRNA.
 SOURCE Bos taurus
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bos.
 REFERENCE 1 Suzuki, M., Morimatsu, M., Yamashita, T., Iwanaga, T. and Syuto, B. A novel serum chitinase that is expressed in bovine liver FEBS Lett. 506 (2), 127-130 (2001)
 MEDLINE 21475601
 REFERENCE 2 (bases 1 to 1529)
 AUTHORS Suzuki, M., Morimatsu, M. and Syuto, B.
 TITLE Direct Submission
 JOURNAL
 FEATURES
 source 1..1529 /organism="Bos taurus" /db_xref="taxon:9913" /tissue_type="liver" 1..1529 /gene="cbp b04" 13..1431 /gene="cbp b04" /codon_start=1 /product="chitin binding protein b04" /protein_id="BAB71805.1" /db_xref="GI:16754895" /translation="MAKILFLTGLAFNLQAQSLQVLCVFSNWAQYRPGLSKPKPD NIDPLCTHLIYAFAGMSNSEITTEWMDVALYSFNDLKKNSQLKLLAIGGNFV TAPFMAPEKNTKPISSVIFKELHGYDGLDFDMEYPGFRGSPQDKHLFTVLVQ ETRAFEEAKQTKNPKLLTAAVAGISNIOAGVEIPOLSOYLDPIHMTVDPHGSM EGYTGNSPLKYPTDGTGNTYLVNVEYAMNWKNGAPAEKLIIGFPAYGHNFLRDA SNNGIAPTSAGGAPGPTREAGFWAYEYICFLKDGATEAWDDSDNQNPYAYKGTETW GYDNNVFRKIAQMLKNNFGAMVWAIDDDFTGTCNKGKPLPLINTLKDALGLKSA TCNASTOSSEPNSESPGNSGSGNSKSSSEGRGYCAGKADGLYPVADNRNAPMNCVNGI TYKQNCITGLVFDTSCHCNWA" 1508..1513 /gene="cbp b04"
 polyA_signal
 BASE COUNT 391 a 405 c 365 g 368 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,17e-140 Length: 1529
 Score: 1991.50 Matches: 348
 Percent Similarity: 87.14% Conservative: 45
 Best Local Similarity: 77.16% Mismatches: 57
 Query Match: 79.88% Indels: 1
 DB: 4 Gaps: 1
 US-10-004-219B-9 (1-452) x AB051629 (1-1529)
 QY 1 TyrAsnLeuIleCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPhe 20
 DB 76 TACCAGCTGGTATGCTACTTCTTAACCTGGGCCAGTACCGTCCAGGCTGGGAAGCTTC 135
 QY 21 LysProAspAspIleAsnProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet 40
 DB 136 AAGCCTGCAACATCGACCCCTGCTCTGCATCTACCTGATCTATGCTTCTGCTGGATG 195

```
QY 41 GlnAsnAsnGluLeuThrThrLeuTrpAsnAspValThrLeuTrpAsnAlaPheAsn 60
DB 196 AGTAACAGTGAGATCACCACCATGGAATGATGTGCTCTCTATAGTTCTTTCAAT 255
QY 61 AspLeuLeuAsnArgAsnSerLeuLeuThrLeuLeuAlaLeuGlyTrpAsnPhe 80
DB 256 GACCTGAAAAAAGAACACCGCAGCTGAAATTTCTCTGCCATTGGAGCGCTGGAACTTT 315
QY 81 GlyThrAlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheIleThr 100
DB 316 GGAACCTGCCCTTCTACTGTATGTTGCCACTCCGGAGAACCGCAAGACTTTCATTTCT 375
QY 101 SerValIleLeuPheLeuArgGlnTrpGlyPheAspGlyLeuAspLeuAspTrpGluTrp 120
DB 376 TCAGTCATCAAAATTTCTACACCATGATGATTTGACGGCTGGATTTGATTTGGAGTAC 435
QY 121 ProGlySerArgGlySerProProGlnAsnAspLeuPheThrValLeuValLeuGlu 140
DB 436 CTTGGCTTTCTGTGGAAGCCCTTCTCAGGACAAAGCATCTCTTCACTGCTTAGTGCAAGAA 495
QY 141 MetArgGluAlaPheGluGlnGlnAlaIleGluSerAsnArgProArgLeuMetValThr 160
DB 496 ACAGTGAGACTTTTGAACAGGAGCAAGCAACCAAGCCAGCGCTGCTGTACAC 555
QY 161 AlaAlaValAlaGlyIleSerAsnIleGlnAlaGlyTrpGluLeuProGluLeuSer 180
DB 556 GCTGCAGTAGCTGCTGGCATCTCTAATATCCAGCGTGGCTATGAGATTTCCCGAGCTATCC 615
QY 181 LysTyrLeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr 200
DB 616 CAGTACCTGGACTTCATCCATGTCATGACCTATGACTTCCATGGCTTCTGGGAAGGCTAC 675
QY 201 ThrGlyGluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeu 220
DB 676 ACAGGAGAGACAGCCCTCTCTACAATATCCCACTGACACTGGCAGCAATACTTACCTC 735
QY 221 AsnValAspTyrValMetAsnTyrTrpLysAsnAsnGlyAlaProAlaGluLysLeuIle 240
DB 736 AATGTGGAATAGCCATGAACACTACTGGAAGAAAATGGGCGCCGCTGAGAAGCTCATC 795
QY 241 ValGlyPheProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIle 260
DB 796 ATTGGATTCCTTCCCTACGAGACAAATTTTCATCTAAGAGACGCTTCCCAACAATGGTATC 855
QY 261 GlyAlaProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrp 280
DB 856 GTGCCCCCATCTCTGAGAGCTGTCTCGGGCCCTACACGAGGAGCTGGGTCTCTGG 915
QY 281 AlaTyrTyrGluLeuCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSer 300
DB 916 GCCTACTATGAGATCTGTGCTTCTCTGAAAGATGGAGCCACTGAGGCATGGGATGACTCT 975
QY 301 GlnGluValProTyrAlaTyrLysAlaAsnGluTrpLeuGlyTyrAspAsnIleLysSer 320
DB 976 CAGAATGTGCCCTATGCCCTACAGGCGCACTGAGTGGTGGTTGATGACAATGTATACAGC 1035
QY 321 PheSerValIleValAlaGlnTrpLeuLysGlnAsnAsnPheGlyValAlaMetIleTrpAla 340
DB 1036 TTCAGAAATCAGGCTCAGTGGCTTAAAGGAGAACAACTTTGGAGGTGCCATGTGCTGGGCC 1095
QY 341 IleAspLeuAspAspPheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSer 360
DB 1096 ATTGACCTGGATTTTACAGGCACTTTCTGCAACAGCGCAAGTTCCCTCTGATAAAT 1155
QY 361 ThrLeuAsnLysAlaLeuGlyIleSerThrGluGlyCysThrAlaProAspValProSer 380
DB 1156 ACCTTGAAGGATGCTCTTGGCTTGAAGAGTGCACCTGTAATGCTCCACGAGCTCCAGT 1215
QY 381 GluProValThrProProGlySerGlySerGlyGlyGlySerSerGlyGlySer 400
DB 1216 GAGCCC---AACAGCAGTCTCTGGAAACGAAGTGGAGGTGGAAATAGAGCAGCAGCTCC 1272
QY 401 GlyGlySerGlyPheCysAlaAspLysAlaAspGlyLeuTyrProValAlaAspAspArg 420
```

```
DB 1273 GAGGCTAGAGGATACTGTGCTGGCAAGCTGATGCCCTTACCCAGTGGCGGATAACAGA 1332
QY 421 AsnAlaPheTrpGlnCysIleAsnGlyIleThrTyrGlnGlnHisCysGlnAlaGlyLeu 440
DB 1333 AATGCTTCTGAACTGCGTAATGGAATCATATCAACACAGATTTGCTGACTGGGCTT 1392
QY 441 ValPheAspThrSerCysAsnCysCysAsnTrp 451
DB 1393 GTCTTTGACACGAGCTGTCTGCTGCAACTGG 1425

RESULT 7
AX405989          1369 bp      DNA      linear      PAT 14-JUN-2002
LOCUS            Sequence 404 from Patent WO0222660.
DEFINITION      AX405989
ACCESSION       AX405989
VERSION         AX405989.1 GI:21439417
KEYWORDS        human.
SOURCE          Homo sapiens
ORGANISM        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE       1
AUTHORS         Tang, Y.T., Liu, C., Zhou, P., Asundi, V., Zhang, J., Zhao, Q.A., Ren, F.,
                Xue, A.J., Yang, Y., Wehrman, T. and Drmanac, R.T.
TITLE           Novel nucleic acids and polypeptides
JOURNAL         Patent: WO 0222660-A 404 21-MAR-2002;
                HYSEQ, INC. (US)
FEATURES        Location/Qualifiers
source          1..1369
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                154..1260
                /note="unnamed protein product"
                /codon_start=1
                /protein_id="CAD35096.1"
                /db_xref="GI:21439418"
                /translation="MVSFPPENRQTFITSVIKFLRQYEPDGLDFDWEYPSGRSGPPQDK
                HLFTVLVQMEAEFQEAQINQRLMVTAAVAGISNIQSGYETPOLSKYLDYTHVM
                TYDLVSGEGVTGENSPLYKYPTDTSNAYLVNVDYVMYKNDGAPAEKLVJGPTYG
                HNFILNSNTGIGNPTSGAGPAGPYAKESGIWAYEYELCTELKNGATGMDAPQSVPY
                AYQNVWGYDNIKSFIDKQWLKNGKFGGAWVAIDLDLDTFTGTCNGKGFPLSTLK
                KALGQSACSTAPAQPIEPTITAAFGSGNGSGSSSSSSGSGSGGSGGFCVANGLYLVAN
                NRNAFHVGVNGVTYQQNCQAGLVFDTSDCCNWA"
BASE COUNT      327 a 393 c 341 g 308 t
ORIGIN
Alignment Scores:
Pred. No.:      1,036-124      Length:      1369
Score:          1778.50      Matches:    315
Percent Similarity: 89.14%      Conservative: 38
Best Local Similarity: 79.55%      Mismatches: 40
Query Match:    71.34%      Indels:      3
DB:              6      Gaps:      1

US-10-004-219B-9 (1-452) x AX405989 (1-1369)
QY 59 PheAsnAspLeuLeuAsnArgAsnSerLysLeuLysThrLeuAlaIleGlyGlyTrp 78
DB 67 TATACCTGAATTTGACGCTCGGAACAGCAGCCAGTCTCTCTGGCCATTGGAGGCTGG 126
QY 79 AsnPheGlyThrAlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPhe 98
DB 127 AACTTCGGGACTGCCCTTCTCACTGCCATGGTTTCTACTCTCGAGAACCGCAGACTTTC 186
QY 99 IleThrSerValIleLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrp 118
DB 187 ATCACTCAGTCATCAAAATTTCTGCGCCAGTATGATGTTTGACGGGCTGCACTTTGACTGG 246
QY 119 GluTyrProGlySerArgGlySerProProGlnAspLysHisLeuPheThrValLeuVal 138
DB 247 GAGTACCTTGGCTCTCGTGGAGGCCCTCTCAGGACAAAGCATCTCTTCACTGCTCTGGTG 306
```



```

Db 451 ATCCATGTATGACCTACACCTCCATGCTCTCTGGAGGCTACACTGGAGAACAGC 510
Qy 206 ProLeuTyrLysThrProThrGluThrGlySerAsnAlaTyrLeuAsnValAspTyrVal 225
Db 511 CCCCTCTACAAATACCCGACTGACCGGACGACCAACCCCTACCTCAATGTGATATGTC 570
Qy 226 MetAsnTyrTyrLysAsnGlyAlaProAlaGluLysLeuIleValGlyPheProGlu 245
Db 571 ATGAACCTATGAGAGCAATGAGACCACTGAGAGCTATCTGTTGGATCCCTACC 630
Qy 246 TyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIleGlyAlaProThrSer 265
Db 631 TATGGACACAACTTCTCTGAGCAACCCCTCCACACTGGAATGGTCCGCCACCTCT 690
Qy 266 GlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTyrAlaTyrGluIle 285
Db 691 GGTGCTGGTCTCTGCTGGGCCCTATGCAAGGAGTCTGGGATCTGGGCTTACTACGAGATC 750
Qy 286 CysThrPheLeuArgSerGlyAlaThrGluValTyrAspAlaSerGlnGluValProTyr 305
Db 751 TGTACCTTCTGAAAATGAGGCACTCAGGATGGGATGCCCTCAGGAAGTGCCTTAT 810
Qy 306 AlaTyrLysAlaAsnGluTyrLeuGlyTyrAspAsnIleLysSerPheSerValLysAla 325
Db 811 GCCTATCAGGCAATGTGTGGTGTGCTATGACACGCTCAAGAGCTTCGATATTAAAGCT 870
Qy 326 GlnTyrLeuLysGlnAsnAsnPheGlyAlaMetIleTyrAlaIleAspLeuAspAsp 345
Db 871 CAATGGCTTAAGCACAACTTTGAGGCGCCATGCTCTGGGCCATTCATCTGGATGAC 930
Qy 346 PheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSerThrLeuAsnLysAla 365
Db 931 TTCACCTGGACCTTCTGCAACCGGCAAGTTCCTTAATCTCCACCTCGAAGAGGCC 990
Qy 366 LeuGlyIleSerThrGluGlyCysThrAlaProAspValProSerGluProValThrThr 385
Db 991 CTTGGCTGCAGAGTCAAGTTGCACGGCTCCAGCTCAGCCCATTCAGGCAATACTGCT 1050
Qy 386 ProPro-----GlySerGlySerGlyGlyGlySerSerGlySerGlyGly 402
Db 1051 GTCCAGTGGCAGCGGAGAACCGGAGCGGAGTAGCAGCTCTGGAGGAGCTCGGGAGGC 1110
Qy 403 SerGlyPheCysAlaAspLysAlaAspGlyLeuTyrProValAlaAspAspArgAsnAla 422
Db 1111 AGTGGATTCTGTGCTGGCAGACCAAGCGCTTACCCCGTGCAATTAACAGAAATGCC 1170
Qy 423 PheTyrGlnCysIleAsnGlyIleThrTyrGlnGlnHisCysGlnAlaGlyLeuValPhe 442
Db 1171 TTCTGGCACTGCGTGAATGGAGTCAGCTACCGACAGCAACTGCCAGCGCGGCTTGCTTC 1230
Qy 443 AspThrSerCysAsnCysCysAsnTyr 451
Db 1231 GACACAGCTGTGATTGCTGCAACTGG 1257

RESULT 9
AB025009
LOCUS
DEFINITION Homo sapiens TSA1902-S mRNA for novel member of chitinase family, complete cds.
VERSION AB025009
KEYWORDS novel member of chitinase family; TSA1902-S.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (sites)
REFERENCE Saito, A., Ozaki, K., Fujiwara, T., Nakamura, Y. and Tanigami, A.
AUTHORS Isolation and mapping of a human lung-specific gene, TSA1902, encoding a novel chitinase family member
JOURNAL Gene, 239 (2), 325-331 (1999)
MEDLINE 20018184

```

```

2 (bases 1 to 1188)
Saito, A., Ozaki, K., Fujiwara, T., Takahashi, E. and Tanigami, A.
Direct Submission
Submitted (16-MAR-1999) Akihiko Saito, Otsuka Pharmaceutical Co., Ltd., Otsuka GEN Research Institute; 463-10 Kagasuno Kawauchi-cho, Tokushima, 771-0192, Japan (E-mail: saito@otsuka.gr.jp, Tel: 81-88-665-2888, Fax: 81-88-637-1035)
FEATURES
    source
        1..1188
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
    gene
        1..1188
            /gene="TSA1902-S"
    CDS
        150..1097
            /gene="TSA1902-S"
            /codon_start=1
            /product="novel member of chitinase family"
            /protein_id="BAA6981.1"
            /db_xref="GI:6467179"
            /translation="MREAFQEAKQIKPRLMTVAAVAAGISNIQSVEIPLSQYLD
            YIHWYDLHGSWEGYTGENSEPLKYPTDTGSNAYLVNDVYMNWYKMGAPAEKLVG
            PFTYGHNFILSNPNTGIGAPTSAGPAGPYAKESGIWAYYEICTFLKNGATQGDAP
            QEVPIAYGNVWVGVDNYSFDIKAKMLKHNKFGGMWALDLDFTGTCFNGQKFPFL
            ISTLKALGLQSGASCTAPAPIEPTITAPSGSGSGSSSGSGSGSGSGSGSGSGSG
            YPVANNRNFVHCVNGVTYQNCQAGLVFDTSCDCNNA"
BASE COUNT 280 a 337 c 309 g 262 t
ORIGIN
Alignment Scores:
Pred. No.: 8,286-100 Length: 1188
Score: 1446.00 Matches: 262
Percent Similarity: 76.35% Conservative: 35
Best Local Similarity: 67.35% Mismatches: 34
Query Match: 59.00% Indels: 59
DB: 9 Gaps: 2

US-10-004-219b-9 (1-452) x AB025009 (1-1188)
Qy 66 AsnSerLysLeuLysThrLeuAlaIleGlyGlyTyrAsnPheGlyThrAlaProPhe 85
Db 91 AACACCCAGCTGAAACCTCTCTGGCCATTGGAGCTGGAACTTCAGGACTGC-CCC--- 146
Qy 86 ThrThrMetValSerThrSerGlnAsnArgGlnThrPheIleThrSerValIleLysPhe 105
Db 146 ----- 146
Qy 106 LeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTyrGluTyrProGlySerArgGly 125
Db 146 ----- 146
Qy 126 SerProGlnAspLysHisLeuPheThrValLeuValLysGluMetArgGluAlaPhe 145
Db 147 -----GAAATGGGTGAAGCTTTT 164
Qy 146 GluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThrAlaAlaValAlaGly 165
Db 165 GAGCAGGAGGCGCAAGCAGATCAACAAAGCCAGGCTGATGGTCACTGCTGCAGTAGCTGT 224
Qy 166 GlyIleSerAsnIleGlnAlaGlyTyrGluIleProGluLeuSerLysTyrLeuAspPhe 185
Db 225 GGCATCTCCAATATCCAGCTGGCTATGAGATCCCACTGTCACAGTACCTGACTGACTAC 284
Qy 186 IleHisValMetThrTyrAspLeuHisGlySerTyrGluGlyTyrThrGlyGluAsnSer 205
Db 285 ATCCATGTCATGACCTACGACCTCCATGGCTCTCTGGAGGGCTACACTGGAGAGACAGC 344
Qy 206 ProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeuAsnValAspTyrVal 225
Db 345 CCCCTCTCAAAATACCCGACTGACCGGACGACCAACCCCTACCTCAATGTGATATGTC 404
Qy 226 MetAsnTyrTyrLysAsnAsnGlyAlaProAlaGluLysLeuIleValGlyPheProGlu 245
Db 405 ATGAACTACTGGAAGGCAATGGAGCACAGCTGAGAAGCTCATCGTTGGATTCCCTACC 464

```

Qy 246 TyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIleGlyAlaProThrSer 265
Db 465 TATGACACAACTTCTATCTGTAGCAACCCCTCCACACTGGATTTGGTGGCCCACTCT 524
Qy 266 GlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrpAlaTyrTyrGluLeu 285
Db 525 GGTGTGTGTCTCTGCTGGCCCTATGCCAAGAGTCTGGGATCTGGGCTTACTACGAGATC 584
Qy 286 CysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSerGlnGluValProTyr 305
Db 585 TGTACTCTCTCGAAATATGAGCCACTAGGATGGATGGCCCTCAGGAATGGCTTAT 644
Qy 306 AlaTyrLysAlaAsnGlnTrpLeuGlyTyrAspAsnIleLysSerPheSerValLysAla 325
Db 645 GCCTATCAGGCAATGTGTGGTGTGGCTATGACAAAGTCAAGTCAAGAGCTTCGATTAAGGCT 704
Qy 326 GlnTrpLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTrpAlaIleAspLeuAspAsp 345
Db 705 CAATGGCTTAAGCAACAACAAATTTGGAGCGCATGGTCTGGGCCATTGATCTGGATGAC 764
Qy 346 PheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSerThrLeuAsnLysAla 365
Db 765 TTCACTGGCARTTCTGCACCAAGGCAAGTTCCCTTAATCTCCACCTCGAAGAGGCC 824
Qy 366 LeuGlyIleSerThrGluGlyCysThrAlaProAspValProSerGluProValThrThr 385
Db 825 CTTGGCTGCAGATGCAAGTTGCAAGGCTCAGCTCAGCCCATTTGAGCCATTAAGTCT 884
Qy 386 ProPro-----GlySerGlySerGlyGlySerGlyGlySerGlySerGlyGly 402
Db 885 GCTCCAGTGGCAGCGGGAACGGAGCGGAGTAGCAGCTCTGGAGGCGAGCTCGGGAGGC 944
Qy 403 SerGlyPheCysAlaAspLysAlaAspGlyLeuTyrProValAlaAspAspArgAsnAla 422
Db 945 AGTGGATCTGTGCTGGCAGACCAAGCGCCCTCTACCCGCGTGCAGAAATACAGAAATGCC 1004
Qy 423 PheTrpGlnCysIleAsnGlyIleThrTyrGlnGlnHisCysGlnAlaGlyLeuValPhe 442
Db 1005 TTCTGGCTGCTGATGATGAGTCACTACGACAGCAGACTGCGAGCGCGGCTTCTTC 1064
Qy 443 AspThrSerCysAsnCysCysAsnTrp 451
Db 1065 GACACAGCTGTGATTTGCTGCAACTGG 1091
RESULT 10
AK098814
LOCUS
DEFINITION
Homo sapiens cDNA FLJ25948 fis, clone STM04207, highly similar to
Homo sapiens acidic mammalian chitinase precursor.
ACCESSION
VERSION
KEYWORDS
SOURCE
Homo sapiens
clone:STM04207.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS
1
Ninomiya, K., Wagatsuma, M., Kanda, K., Kondo, H., Yokoi, T.,
Kodaira, H., Furuya, T., Takahashi, M., Kikkawa, E., Omura, Y., Abe, K.,
Kamihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M.,
Suzuki, Y., Hata, H., Nakagawa, K., Mizuno, S., Morinaga, M.,
Kawamura, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H.,
Nishikawa, T., Sugiyama, A., Kawakami, B., Nagai, K., Isogai, T. and
Sugano, S.
NEDO human cDNA sequencing project
TITLE
JOURNAL
REFERENCE
2 (bases 1 to 4250)
Sugano, S. and Suzuki, Y.
Direct Submission
TITLE
JOURNAL
Submitted (08-JUL-2002) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure, Human Genome

Center; Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
(E-mail:cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction and 5'-end one pass sequencing: Institute of Medical
Science, University of Tokyo, Laboratory of Genome Structure, Human
Genome Center; 3'-end one pass sequencing: RAB; clone selection for
full insert sequencing: RAB and Helix Research Institute.

FEATURES
source

1..4250
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="STM04207"
/tissue_type="stomach mucosa"
/clone_lib="STM"
/note="cloning vector: pME18SFL3"

BASE COUNT 1262 a 868 c 966 g 1154 t
ORIGIN

Alignment Scores:
Pred. No.: 8,27e-97 Length: 4250
Score: 1414.50 Matches: 249
Percent Similarity: 87.69% Conservative: 36
Best Local Similarity: 76.62% Mismatches: 37
Query Match: 56.74% Indels: 3
DB: 9 Gaps: 1

US-10-004-219B-9 (1-452) x AK098814 (1-4250)

Qy 130 AspyLysHisLeuPheThrValLeuValLysGluMetArgGluAlaPheGluGlnGluAla 149
Db 3174 GATGTCCTATTATGCTTATTTCTGTAGGAATGCTGAGCTTTTGAGCAGAGGCC 3233
Qy 150 IleGluSerAsnArgProArgLeuMetValThrAlaAlaValAlaGlyGlyIleSerAsn 169
Db 3234 AAGCAGATCAACAGCCCGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3293
Qy 170 IleGlnAlaGlyTyrGluLeuProGluLeuSerLysTyrLeuAspPheIleHisValMet 189
Db 3294 ATCCAGTCTGGCTATGAGATCCCCCACTGACAGTACCTGGAGTATGATGATGATGATG 3353
Qy 190 ThrTyrAspLeuHisGlySerTyrGluGlyTyrThrGlyGluAsnSerProLeuTyrLys 209
Db 3354 ACCTACGACCTCCATGGCTCTCTGGAGGGCTACACTGGAGAGAACAGCCCCCTTACAAA 3413
Qy 210 TyrProThrGluThrGlySerAsnAlaTyrLeuAsnValAspTyrValMetAsnTyrTrp 229
Db 3414 TACCGAGTCTGACCGGAGCAACGCTTACCTCAATGTGGATTATGTCATGAATCTACTGG 3473
Qy 230 LysAsnAsnGlyAlaProAlaGluLysLeuIleValGlyPheProGluTyrGlyHisThr 249
Db 3474 AAGGCAATGGAGCACCAGCTGAGAAGTCTATCGTTGGATTCCCTACCTATGGACACAA 3533
Qy 250 PheIleLeuArgAsnProSerAspAsnGlyIleGlyAlaProThrSerGlyAspGlyPro 269
Db 3534 TTCATCTGAGCAACCCCTCCACACTGGATTGGTGGCCCACTCTGGTCTGGTCTCT 3593
Qy 270 AlaGlyAlaTyrThrArgGlnAlaGlyPheTrpAlaTyrTyrGluIleCysThrPheLeu 289
Db 3594 GCTGGGCCCTATGGCAAGAGTCTGGGCTTACTACGAGATCTGTACCTTCTCTG 3653
Qy 290 ArgSerGlyAlaThrGluValTrpAspAlaSerGlnGluValProTyrAlaTyrLysAla 309
Db 3654 AAAATGGAGCCACTCAGGGATGGGATGCCCTCAGGAAGTGCCTTATGCTTATCGAGGC 3713
Qy 310 AsnGluTrpLeuGlyTyrAspAsnIleLysSerPheSerValLysAlaGlnTrpLeuLys 329
Db 3714 AATGTGGTGGTGGCTATGACAACTCAAGAGCTTCGATATTAGGCTCAATGGCTTAA 3773
Qy 330 GlnAsnAsnPheGlyGlyAlaMetIleTrpAlaIleAspLeuAspPheThrGlySer 349

```

Db      3774  CACAACAATTTGGAGCGCCATGTCTGGGCCATTGATCTGGATGACTTCACTGGCACT 3833
Qy      350  PheCysAspGlnGlySerPheProLeuThrSerThrLeuAsnLysAlaLeuGlyIleSer 369
Db      3834  TTCTGACACAGCGGCAAGTTTCCCTTAATCTCCACCTGAAGAGGCCCTTCGGCCTGCGAG 3893
Qy      370  ThrGluGlyCysThrAlaProAspValProSerGluProValThrThrProPro----- 387
Db      3894  AGTGCAAGTTGACAGGCTCAGCTCAGCCATTGAGCCATAACTGCTGCTCCAGTGCG 3953
Qy      388  ---GlySerGlySerGlyGlySerSerGlySerSerGlySerGlySerGlyPheCys 406
Db      3954  AGCGGAAGCGGAGCGGAGTAGCAGCTCTGGAGGAGCTCGGGAGCGAGTGGATTCTGT 4013
Qy      407  AlaAspLysAlaAspGlyLeuThrProValAlaAspAspArgAsnAlaPheTrpGlnCys 426
Db      4014  GCTGTCAGAGCCCAACGGCTCTACCCGCGGCAATAAAGAAATGCCCTTCGGCACTGC 4073
Qy      427  IleAsnGlyIleThrTyrrGlnGlnHisCysGlnAlaGlyLeuValPheAspThrSerCys 446
Db      4074  GTGAATGAGTACAGTACCGACGACGAGTCCAGCGCGGCTTGTCTTCGACACCACTGT 4133
Qy      447  AsnCysCysAsnTrp 451
Db      4134  GATTGCTGCAACTGG 4148

RESULT 11
M94584
LOCUS      M94584      1525 bp      mRNA      linear      ROD 15-JUN-2001
DEFINITION Mus musculus secretory protein precursor (Yml) mRNA, complete cds.
ACCESSION M94584
VERSION   M94584.2 GI:11140876
KEYWORDS
SOURCE    Mus musculus.
ORGANISM  Mus musculus
REFERENCE 1 (bases 1 to 1525)
AUTHORS   Chang,N.-C.A., Liu,C.-H. and Chang,A.C.
TITLE     Molecular characterization of a secretory protein (YM-1)
          transiently expressed by activated murine peritoneal macrophages
JOURNAL   Unpublished
REFERENCE 2 (bases 1 to 1525)
AUTHORS   Chang,N.C., Hung,S.I., Hwa,K.Y., Kato,I., Chen,J.E., Liu,C.H. and
          Chang,A.C.
TITLE     A macrophage protein, Yml, transiently expressed during
          inflammation is a novel mammalian lectin
JOURNAL   J. Biol. Chem. 276 (20), 17497-17506 (2001)
MEDLINE   21264517
PUBMED    11297523
REFERENCE 3 (bases 1 to 1525)
AUTHORS   Chang,N.-C.A., Liu,C.-H. and Chang,A.C.
TITLE     Direct Submission
          Submitted (27-APR-1993) Microbiology and Immunology, National
          Yang-Ming University, 155, Sect. II, Li-Noon St, Taipei, Taiwan
          112, R.O.C.
REFERENCE 4 (bases 1 to 1525)
AUTHORS   Chang,N.-C.A., Liu,C.-H. and Chang,A.C.
TITLE     Direct Submission
          Submitted (13-NOV-2000) Microbiology and Immunology, National
          Yang-Ming University, 155, Sect. II, Li-Noon St, Taipei, Taiwan
          112, R.O.C.
REMARK    Sequence update by submitter.
COMMENT   On Nov 13, 2000 this sequence version replaced gi:202441.
FEATURES
         Location/Qualifiers
            1..1525
                /organism="Mus musculus"
                /db_xref="taxon:10090"
                /clone_lib="PEC in lambda gt11 and ZAPII"
                /dev_stage="adult"
            1..1525
                /gene="Yml"
            13..1209
                /CDS

gene
CDS

```

```

/gene="Yml"
/codon_start=1
/product="secretory protein precursor"
/protein_id="AAB62394.2"
/db_xref="GI:11140877"
/translation="MAKILVTGLALILNVLGSSYQLMCMYTTWAKDRPIEGSPKPG
NIDPLCTHLIYAFAGMNNETITHTQDLRYEALNGKDKNTLKTLLAIGWKFV
PAPFSAMVSTPNQRIQIFQSVIRFLQYDFGLNDMQLDQYSGRSPKDKHLFSLVK
EMRKAFEBESVEKDIPLLSTGAGIIDVKSGYKI PELSQSLDYIOVMTYDLHDPK
DGYTGNSPLKSPYDVGKSDANLVDISIYMKDHGASEKLIYGFPRAYGHTFLLSDP
SKTGIGATTISTGPPKKTDESGLLAYIEVCTFLNEGATEVNDAPQEVPIAYQGNWV
GYNVRSPLKKAQWLKONNLGSAVWPLMDMDFSGSFQHRHFLTSTLKGDLNLSA
SCKGFP"
sig_peptide 13..75
mat_peptide 76..1206
polyA_signal 1487..1492
polyA_site 1507
BASE COUNT 434 a 343 c 335 g 413 t
ORIGIN
Alignment Scores:
Pred. No.: 2,98e-95 Length: 1525
Score: 1387.00 Matches: 249
Percent Similarity: 80.8% Conservative: 55
Best Local Similarity: 66.2% Mismatches: 72
Query Match: 55.64% Indels: 0
DB: 10 Gaps: 0

US-10-004-219B-9 (1-452) x M94584 (1-1525)
Qy      1  TyrAsnLeuIleCysTyrrPheThrAsnTrpAlaGlnTyrrArgProGlyLeuGlySerPhe 20
Db      76  TACCAGCTGATGTGCTACTATACCACTGGGCTAAGGACAGGCCCAATAGAAGCGAGTTTC 135
Qy      21  LysProAspAspIleAsnProCysLeuCysThrHisLeuIleTyrrAlaPheAlaGlyMet 40
Db      136  AAACCTGGTATATATTGACCCCTGCTGTACTACCTACCTGATCTATGCTTGTCTGGAATG 195
Qy      41  GlnAsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrrLysAlaPheAsn 60
Db      196  CAGATATATGAGTACACTACACATGACGACAGACTTGGCTGACTATGAGCATTTGAT 255
Qy      61  AspLeuLysAsnArgAsnSerLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPhe 80
Db      256  GGTCTGAAAGACAGACACACTGAGCTAAATACTCTCTGGCCATTGGAGGATGGAAGTTT 315
Qy      81  GlyThrAlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheIleThr 100
Db      316  GGACCTGCCCCGTTTCAGTGCCATGCTCTACTCTCTCAGAACCGTCAGATATTTCAAT 375
Qy      101  SerValIleLysPheLeuArgGlnTyrrGlyPheAspGlyLeuAspLeuAspTrpGluTyrr 120
Db      376  TCAGTTATCAGATTCCTTCGTCATATATACTTTGATGGCTCAACCTGAGCTGGCAGTAC 435
Qy      121  ProGlySerArgGlySerProGlnAspLysHisLeuPheThrValLeuValLysGlu 140
Db      436  CTGGGTCTCGAGGAAGCCCTCTTAAGGACAAACATCTCTTCAGTGTCTGCTGGAAGAA 495
Qy      141  MetArgGluAlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThr 160
Db      496  ATGCGTAAAGCTTTTGAAGGAAGAATCTCTGGAGAAAGACATTTCCAAAGGCTGCTACT 555
Qy      161  AlaAlaValAlaGlyIleSerAsnIleGlnAlaGlyTyrrGluIleProGluLeuSer 180
Db      556  TCCACAGGAGCAGGAGNATCATTTGACGATATCAAGTCTGGGTACAGATCCCTGAATGCT 615
Qy      181  LysTyrrLeuAspPheIleHisValMetThrTyrrAspLeuHisGlySerTrpGluGlyTyrr 200
Db      939  :

```

Db 616 CAGTCTCTTACTATATTCAGGTCATGACATATGATCTCCATGATCTTAAGGATGCTAC 675
Qy 201 ThrGlyGluAsnSerProLeuTyrIlystYrProThrGluThrGlySerAsnAlaTyrLeu 220
Db 676 ACTGAGAAATAGTCCCTCTATAATCTCATATGATGATGGAAGAGTCTGATCTC 735
Qy 221 AsnValAspTyrValMetAsnTyrTrpIlyAsnAsnGlyAlaProAlaGluIlyLeu 240
Db 736 AATGTGGATTCAATCATTTCTTACTTGAAGGACCATGAGGAGCTTCTGAGAAGCTCAT 795
Qy 241 ValGlyPheProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIle 260
Db 796 GTGGGATTTCCAGCATATGCGCATACCTTATCTCTGAGTGACCTTCTAAGACTGGAAT 855
Qy 261 GlyAlaProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrp 280
Db 856 GTGGCCCTTACAACTAGTACTGCGCCACAGAAAGTACACAGATGAATCAGGACTCTCG 915
Qy 281 AlaTyrTyrGluIleCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSer 300
Db 916 GCTTACTATGAGGTTGTACATTTCTTGAATGAAGGAGCCACTGAGGCTCGGATGCCCC 975
Qy 301 GlnGluValProTyrAlaTyrIlyAsnAlaGlnGluTrpLeuGlyTyrAspAsnIlySer 320
Db 976 CAGGAAGTACCTTATGCTTATCAGGTAATGAGTGGTGGTATGACATGTCAGGAGC 1035
Qy 321 PheSerValIlyAlaGlnTrpLeuIlyGlnAsnAsnPheGlyGlyAlaMetIleTrpAla 340
Db 1036 TTCAGTTGAAGGCTCAGTGGCTCAAGGACACAAATTTAGGAGGTGCCGTGTGGCC 1095
Qy 341 IleAspLeuAspAspPheThrGlySerPheCysAspGlnGlyIlySerPheProLeuThrSer 360
Db 1096 CTGGACATGATGATCTCAGTGGTCTTCTGTGACCAGAGACATTTCCCTCTGACATCT 1155
Qy 361 ThrLeuAsnAlaLeuGlyIlySerThrGluGlyCysThrAlaPro 376
Db 1156 ACTTTAAGGAGATCTCAATATACACAGTGAAGTTGAAGGGCCCT 1203
RESULT 12
D87757 1506 bp mRNA linear ROD 12-JAN-2000
LOCUS Mus musculus mRNA for ECF-L precursor, complete cds.
DEFINITION D87757
ACCESSION D87757.2 GI:6015435
VERSION ECF-L precursor.
KEYWORDS Mus musculus bone marrow cDNA to mRNA.
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1506)
AUTHORS Owashi, M., Arita, H. and Hayai, N.
TITLE Identification of a novel eosinophil chemotactic cytokine (ECF-L) as a chitinase family protein
JOURNAL J. Biol. Chem. 275 (2), 1279-1286 (2000)
MEDLINE 20092902
REFERENCE 2 (bases 1 to 1506)
AUTHORS Owashi, M.
TITLE Direct Submission
JOURNAL Submitted (06-SEP-1996) Makoto Owashi, The University of Tokushima, Faculty of Integrated Arts and Sciences; 1-1 Minami-Johananjima, Tokushima, Tokushima 770, Japan (E-mail:owashi@ias.tokushima-u.ac.jp, Tel:0886-56-7261, Fax:0886-56-7298)
COMMENT On Oct 7, 1999 this sequence version replaced gi:1545818.
FEATURES
source Location/Qualifiers
1..1506
/organism="Mus musculus"
/db_xref="taxon:10090"
/tissue_type="bone marrow"
1..1506
/gene="mECF-L"
1..1197

/gene="mECF-L"
/codon_start=1
/product="ECF-L precursor"
/protein_id="BAA13458.2"
/db_xref="GI:6015435"
/translation="MAKULIVTGLAILNLVOLGSSYQLMCYTTSWAKDRPIEGSPKPG
NIDLCPLHLIYAFAGMNNIYTHQDLRDYEALNGDKNTKLTLLAIGGKFKG
PASFSAMVTPQNRQIFQISVIRFLQYFNGLNDQYPSGRSGSPDKHLFSLVK
EMKAFESVEKDI PRLLTSTGAGI DVIKSGYKI PELSQSLDYIQVMTYDLHDPK
DGYTGENSPKSVYDICKSADLNVDSIISYKDHGAASEKLIVGPPAVGHTFLLSDP
SKTGIGATISTGPPCKTDESGLLAIYYEVECTFLNEGATEWDARQEPVAYAQNEW
GYDNVRSFKAKMLKONLGVVWPLMDDDFSGSFQHRHFPLTSLTKGLNLHSA
SCKGPY"
BASE COUNT 423 a 338 c 333 g 412 t
ORIGIN
Alignment Scores:
Pred. No.: 1,176-94 Length: 1506
Score: 1379.00 Matches: 248
Percent Similarity: 80.59% Conservative: 55
Best Local Similarity: 65.96% Mismatches: 73
Query Match: 55.31% Indels: 0
DB: 10 Gaps: 0
US-10-004-219B-9 (1-452) x D87757 (1-1506)
Qy 1 TyrAsnLeuIleCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPhe 20
Db 64 TACCAGCTGATGTGCTACTATACCAGTTGGGCTAAGGACAGGCCCAATAGAGGGAGTTTC 123
Qy 21 LysProAspAspIleAsnProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet 40
Db 124 AAACCTGGTAATATGACCCCTGCTGTACTACCTGATCTATGCTTTCTGGAATG 183
Qy 41 GlnAsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrIlyAsnAlaPheAsn 60
Db 184 CAGAATAATGAGATCACTTACACATGACGACGAGACTTGGCTGATGATGAATGAAT 243
Qy 61 AspLeuIlyAsnAsnArgAsnSerIlyLeuIlyThrLeuAlaIleGlyGlyTrpAsnPhe 80
Db 244 GGTCTGAAGACACAGAACACTGAGCTAAACATCTCTGGCCATTTGGAGGATGGAAGTTT 303
Qy 81 GlyThrAlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheIleThr 100
Db 304 GGACCTGCTCGTTCAGTGGCCATGCTCTACTCTCTCAGAACCGTCAGATATTCATTAG 363
Qy 101 SerValIleIlyPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyr 120
Db 364 TCAGTTATCAGATTCCTTCGTCATATAAATCTTGTATGGCTCAACCTGGACCTGGCAGTAC 423
Qy 121 ProGlySerArgGlySerProGlnAspIlyHisLeuPheThrValLeuValIlyGlu 140
Db 424 CTGGGTCTCGAGGAGGCCCTCTAAGGACAAACATCTCTCAGTGTCTGTGGTGAAGGAA 483
Qy 141 MetArgGluAlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThr 160
Db 484 ATGCGTAAAGCTTTTGGAGGAGAACTCTGGAGAGAGACATTCCAAGGCTGCTACTACT 543
Qy 161 AlaAlaValAlaGlyIlySerAsnIleGlnAlaGlyTyrGluIleProGluLeuSer 180
Db 544 TCCACAGGAGGAGGAATCATTCAGCTTAATCAAGTCTGGGTACAAAGATCCCTGAACCTGT 603
Qy 181 LysTyrLeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr 200
Db 604 CAGTCTCTGATATATTCAGGTCAATGATATGATCTCCATGATCCTTAAGGATGGCTAC 663
Qy 201 ThrGlyGluAsnSerProLeuTyrIlySerTyrProThrGlyThrGlySerAsnAlaTyrLeu 220
Db 664 ACTGGAGAAATAGTCCCTCTATAAATCTCATATGATGACATTCGAAAGAGTCTGATCTC 723
Qy 221 AsnValAspTyrValMetAsnTyrTrpIlyAsnAsnGlyAlaProAlaGluIlyLeu 240
Db 724 AATGTGGATTCAATCATTTCTTACTTGAAGGACCATGAGGAGCTTCTGAGAAGCTCAT 783

QY 241 ValGlyPheProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIle 260
 Db 784 GTGGGATTTCCAGCATATGGGCATACCTTTATCCTGAGTGACCTTCTAAGACTGGAAIT 843
 QY 261 GlyAlaProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrp 280
 Db 844 GTGCCCCCTACAAATAGTAGTCTGGCCCCCAGGAGGAAAGTACACAGATGAATCAGGACTCCTG 903
 QY 281 AlaTyrTyrGluLeuCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSer 300
 Db 904 GCTTACTATGAGTTTGTACATCTTCTGAATGAAGGAGCCACTGAGGTCTGGGATGCCCC 963
 QY 301 GlnGluValProTyrAlaTyrLysAlaAsnGluTrpLeuGlyTyrAspAsnIleLysSer 320
 Db 964 CAGGAAGTACCCTATGCCCTATCAGGTGTAATGAGTGGTGGTTATGACATGTGAGAGC 1023
 QY 321 PheSerValLysAlaGlnTrpLeuLysGlnAsnAspPheGlyAlaMetIleTrpAla 340
 Db 1024 TTCAAGTTGAAGGCTCAGTGGCTCAAGGACAAATTTAGGAGGTGCCGTGGTGGCCC 1083
 QY 341 IleAspLeuAspAspPheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSer 360
 Db 1084 CTGGACATGATGACTTCAGTGGTCTTCTGTCCACGAGACATTTCCCTCTGCATCT 1143
 QY 361 ThrLeuAsnLysAlaLeuGlyIleSerThrGluGlyCysThrAlaPro 376
 Db 1144 ACTTTAAGGGAGATCTCAATATACACAGTGCAGTTCGAAGGGCCCT 1191

RESULT 13
 AY049765
 LOCUS
 DEFINITION Mus musculus secreted protein precursor Ym2 mRNA, complete cds.
 ACCESSION AY049765
 VERSION AY049765.2 GI:22123906
 KEYWORDS
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 1209)
 Webb,D.C., McKenzie,A.N. and Foster,P.S.
 Expression of the Ym2 lectin-binding protein is dependent on
 interleukin (IL)-4 and IL-13 signal transduction: identification of
 a novel allergy-associated protein
 J. Biol. Chem. 276 (45), 41969-41976 (2001)
 21551268
 11553626
 2 (bases 1 to 1209)
 Webb,D.C. and Foster,P.S.
 Direct Submission
 Submitted (02-AUG-2001) Biochemistry and Molecular Biology, John
 Curtin School of Medical Research, Australian National University,
 Off Mills Rd, Acton, Canberra, ACT 2601, Australia
 3 (bases 1 to 1209)
 Webb,D.C. and Foster,P.S.
 Direct Submission
 Submitted (06-AUG-2002) Biochemistry and Molecular Biology, John
 Curtin School of Medical Research, Australian National University,
 Off Mills Rd, Acton, Canberra, ACT 2601, Australia
 Sequence update by submitter
 On Aug 6, 2002 this sequence version replaced gi:15705156.
 Location/Qualifiers
 1..1209
 /organism="Mus musculus"
 /strain="BALB/c"
 /db_xref="taxon:10090"
 /note="isolated from the lung lavage from allergic mice"
 1..1209
 /note="expression regulated by IL-4 and IL-13 signaling
 via the IL-4R alpha subunit"
 /codon_start=1
 /product="secreted protein precursor Ym2"

CDS

/protein id="AAL03953.2"
 /db_xref="GI:22123907"
 /translation="MAKILVLTGLAILNLVOLGSSYQLMVCYVTSWAKDRPTGSGPKG
 NIDPCLTHLIYAFAGMKNNEITYLSEQLDRYDALAGLKDRLTELKTLALGGWKFG
 PAPSSVSTPQNRQTFIKSVIRFLRQYNFDGLNDWQYPGSRGSPDKKHLFSLVQ
 EMRAFEESTLNHPRLLSTGAGFIDVIKSGYKIPELSQSLDIYQVMTYDLHDPK
 NGYTGENSPYSPYDIKSDADLVDSITTYKDHGASERKLVGFPPAHTFILSDP
 SKNGIDPDTVSAGPGKYTNEQGLLAVPEICTFLNEGATEIFDQTEVPYVAYLGNWV
 CYDNVSPFLKAQWLKONNLGAVVWPLDMDDFSSGFCHQGRFFPLTTTLKRDNLNVHSA
 SCKASYRGL"

sig_peptide 1..63
 BASE COUNT 326 a 273 c 282 g 328 t
 ORIGIN

Alignment Scores: 2.57e-94 Length: 1209
 Pred. No.: 1373.00 Matches: 245
 Score: 81.60% Conservative: 61
 Best Local Similarity: 65.33% Mismatches: 69
 Query Match: 55.07% Indels: 0
 DB: 10 Gaps: 0

US-10-004-219B-9 (1-452) x AY049765 (1-1209)

QY 1 TyrAsnLeuIleCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPhe 20
 Db 64 TACCAGCTGATGCTGCTACTATACAGCTGGGCTAAGGACAGGCCAACAGAGGGAGTTTC 123
 QY 21 LysProAspAspIleAsnProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet 40
 Db 124 AAACCTGTAATATTGACCCCTGCTGTGTACTACCTGATCTATGCTGCTGGGATG 183
 QY 41 GlnAsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrLysAlaPheAsn 60
 Db 184 AAGAATAATGAGATCACCTTACTTAAGTGAGCAAGACTTGGCGTACTATGACGATTAAT 243
 QY 61 AspLeuLysAsnArgAsnSerLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPhe 80
 Db 244 GGTCTGAAGACAGGACACACTGAGCTAAACACTCTCTGGCCATTGGAGGATGGAAGTTT 303
 QY 81 GlyThrAlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheIleThr 100
 Db 304 GGACCTGCGCCGCTTCAGTTCCTGCTCTCTCTCCTCAGAACCCGTCAGACATTCATTAA 363
 QY 101 SerValIleLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyr 120
 Db 364 TCAGTTATCAGATTCCTTCGTCATATTAATCTTGAAGCCCTCAACCTGGACTGGCGATAC 423
 QY 121 ProGlySerArgGlySerProGlnAspLysHisLeuPheThrValLeuValLysGlu 140
 Db 424 CCTGGGTCTCGAGGAGGCCCTCTTAAGGACAAACATCTCTTCAGTGTCTGGTGAGGAA 483
 QY 141 MetArgGluAlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThr 160
 Db 484 ATGCGTAAAGCTTTTGAAGGAGAAATCCACTTTGAACCAACATTCACCAAGCTGCTACT 543
 QY 161 AlaAlaValAlaGlyGlyIleSerAsnIleGlnAlaGlyTyrGluIleProGluLeuSer 180
 Db 544 TCCACAGAGCTGGATTCATTGACGTAATCAAGTCTGGGTACAGATCCCTGAATGCT 603
 QY 181 LysTyrLeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr 200
 Db 604 CAGTCTCTCGACTATATTTCAGTCAATGATGATGATGATGATGATGATGATGATGATGAT 663
 QY 201 ThrGlyGluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeu 220
 Db 664 ACTGGAGAAAATAGTCCCTCTATAAATCTCCATATGATGATGATGATGATGATGATGAT 723
 QY 221 AsnValAspTyrValMetAsnTyrTrpLysAsnGlnGlyAlaProAlaGluLysLeuIle 240
 Db 724 AATGTGGATTCATATTACTTACTTGGAGGAGGACCATGAGGAGGAGGAGGAGGAGGAGG 783
 QY 241 ValGlyPheProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIle 260


```
Db 784 GTGGGATTTCAGCATATGGTCATACCTTTATCTCGAGTGACCCCTTCTAAGAAATGGAATA 843
Qy 261 GlyAlaProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGluAlaGlyPheTrp 280
Db 844 GTGACCTTACCTGTTAGTCTGCGACACACAGAAAGTACAAAGTGAACAGAGACTCTCG 903
Qy 281 AlaTyrTyrGluLeuLeuGlyCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSer 300
Db 904 GCTTACTTTGAGATTGTGACATTTCTGAATGAAGAGAGCCACTGAGATCTTTGATGCCACC 963
Qy 301 GlnGluValProTyrAlaTyrIysAlaGlnGluTrpLeuGlyTyrAspAsnIleIysSer 320
Db 964 CAGGAAGTACCTATGCTGCTATCTGGTATGATGAGTGGTGGTTGTTATGACAAATCTCAGGAGC 1023
Qy 321 PheSerValIysAlaGlnTrpLeuIysGlnAsnAsnPheGlyGlyAlaMetIleTrpAla 340
Db 1024 TTCAGTTGAAGCTCAGTGGCTCAGGACAAATTTAGGAGGTCCTGTCCTGGGCC 1083
Qy 341 IleAspLeuAspAspPheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSer 360
Db 1084 CTGGACATGGATGACTTCAGTGGTTCTTCTCTCACCAGGAGCTTTCCTCTGACAAC 1143
Qy 361 ThrLeuAsnLysAlaLeuGlyIleSerThrGluGlyCysThrAla 375
Db 1144 ACTTTAAAGAGAGATCTGTAATGTACACAGTGCAGAGTTGCAAGGCC 1188

RESULT 14
BJA345054 1562 bp mRNA linear VRT 14-JUN-2002
LOCUS BJA345054
DEFINITION Bufo japonicus mRNA for toad pancreatic chitinase (tpCase gene).
ACCESSION AJ345054
VERSION AJ345054.1 GI:15963340
KEYWORDS toad pancreatic chitinase; tpCase gene.
SOURCE Japanese toad.
ORGANISM Bufo japonicus

REFERENCE
1 Oshima, H. Department of Department of Biology, Waseda University, Tokyo, Japan
2 Oshima, H., Miyazaki, R., Ohe, Y., Hayashi, H., Kawamura, K. and Kikuyama, S. Isolation and sequence of a novel amphibian pancreatic chitinase Comp. Biochem. Physiol. 132 (2), 381-388 (2002)
3 Oshima, H. Direct Submission
JOURNAL Submitted (03-SEP-2001) Oshima H., Department of Biology, Waseda University, Nishiwaseda 1-6-1, Shinjuku-ku, Tokyo, 169-8050, JAPAN
FEATURES
Location/Qualifiers
1..1562
/organism="Bufo japonicus"
/db_xref="taxon:8387"
/tissue_type="pancreas"
/country="Japan"
1..1562
/gene="tpCase"
2..1468
/gene="tpCase"
/EC_number="3.2.1.14"
/function="chitin fragmentation"
/codon_start=1
/evidence="experimental"
/product="toad pancreatic chitinase"
/protein_id="CAC87888.1"
/db_xref="GI:15963341"
/db_xref="STREMBL:Q90W34"
/translation="MLLWAGLEFLLHLVQLGSGTYKLVVCYFTNWSQYRPDQKYPGNID
POLCTHLVYAFATMNEHKIAPYEWNDVLYKQFNDLKQKNKLVLLAIGWNFTQK
```

```
FTDMVASGNRSIFIKSVIAYLRQNNFDGIDLDFFEYSGSRGSPBEDKQRFVLTQEML
DAFNEEARSSGLPRLLITAAVSAAGKGTIDAGYEIAKIQQLLDFISVMTYDFHGGWDTQ
SGHNSPLCKSGTDYGLQYFNFIHEAMNWKNGGAPAEKLLGLLFTYGTFRPNPNMC
DVGIPVSGAGSAGPYTREAGWAVEYELCTWLSGSTVKWI PQRPVPYACKSNWVGFDP
QESVECKVRLKESGFGGMWVAIDLDFFGRPCNOCGRYPLINHLSKSLLEGSTVNCPC
EICGSIPTPTPSATTTTTTTAKDCCTTPPEPPTPPVPVVIDVDNFCVEKTDGLH
VNPLNTNKFYICANGRTYSMKCADGLVFQASNCNCNP"
2..55
/gene="tpCase"
/evidence="experimental"
56..1465
/mat_peptide
/gene="tpCase"
/product="toad pancreatic chitinase"
/EC_number="3.2.1.14"
/evidence="experimental"
BASE COUNT 470 a 329 c 345 g 418 t
ORIGIN
Alignment Scores:
Pred. No.: 4,87e-94 Length: 1562
Score: 1371.00 Matches: 259
Percent Similarity: 65.42% Conservative: 55
Best Local Similarity: 53.96% Mismatches: 128
Query Match: 54.99% Indels: 38
DB: 5 Gaps: 6
US-10-004-219B-9 (1-452) x BJA345054 (1-1562)
Qy 1 TyrAsnLeuIleCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPhe 20
Db 56 TACAAACTGGTGTGTACTTTTACAAACTGGTCACAGTATCGCCCTGACGAGGAAGTAT 115
Qy 21 LysProAspAspIleAsnProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet 40
Db 116 GTGCCAGGGAATATCGACCCGAGCTCTGCACACATCTCGTATATGCTTCGCCACCATG 175
Qy 41 GlnAsnAsnGluIleThrIleGluTrpAsnAspValThrLeuTyrLysAlaPheAsn 60
Db 176 AATGAGCATAGATGCGCCATATGAAATGGAATGATGATGCTTTATACAAACAGTTTAA 235
Qy 61 AspLeuLysAsnArgAsnSerLysLeuLysThrLeuAlaIleGlyLysTrpAsnPhe 80
Db 236 GACCTAAAGCAGAAAATAAAACCCCTGGTGACATTTATTGGCAATCGGTGGTGAACATT 295
Qy 81 GlyThrAlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheIleThr 100
Db 296 GGAACACAGAGATTTCACAGATATGTTCTCTCTGGAACCCGTCATTTTCATTAA 355
Qy 101 SerValIleLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyr 120
Db 356 TCTGTGATTGCATATCTACGACAAATAATTTTGTATGGAATAGACCTTGATTTGAATAT 415
Qy 121 ProGlySerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValLysGlu 140
Db 416 CCAGGATCTCGAGGAAGTCTCCAGAGGATAAGCAGCATTTACTGTACTGATACAGAA 475
Qy 141 MetArgGluAlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThr 160
Db 476 ATGCTTGATGATTTAATGAAGAAGCAAGATCTCTCGGTCTTCCTCGACTCTCTAATCA 535
Qy 161 AlaAlaValAlaGlyGlyIleSerAsnIleGlnAlaGlyTyrGluIleProGluLeuSer 180
Db 536 GCAGCAGTGTCCAGCAGAAAAGGCCACCATCGATGCTGATATGATAGATAGCAAGATCG 595
Qy 181 LysTyrLeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr 200
Db 596 CAATTGCTGGATTTTATCAGTGTGATGATCTTATGATTTTCCATGGTGGATGGACACTCA 655
Qy 201 ThrGlyGluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeu 220
Db 656 TCTGGACACACAGCCCTCTGTGCAAGAGTTCCACTGACTACGAGATTTGCAACTTTT 715
Qy 221 AsnValAspTyrValMetAsnTyrTrpLysAsnAsnGlyAlaProAlaGluLysLeuIle 240
```

Db 716 AACATTCATTTTCCCATGAATATTGGAATAAACAAGGGGCCCTCGAGAGAACTCCTT 775
Qy 241 ValGlyPheProGluTyrGlyHisThrPheLeuArgAsnProSerAspAsn----- 258
Db 776 CTTGGGATTTCCCATCATATGCGCAACATTC-----AGAAACCAATCTTACATGTGT 829
Qy 259 GlyIleGlyAlaProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGly 278
Db 830 GATGTGGGATCCCGTGTCTGGAGAGATCAGCTGGGCCATACACAGAGAGGTGGC 889
Qy 279 PheTrpAlaTyrTyrGluLeuGlyCysThrPheLeuArgSerGlyAlaThrGluValTrpAsp 298
Db 890 TTTCTGGGCTTACTATGAGATCTGCATCTGGCTT---TCAGGAGACACAGTGAATGGATT 946
Qy 299 AlaSerGlnGluValProTyrAlaTyrLysAlaAsnGluTrpLeuGlyTyrAspAsnIle 318
Db 947 CCAGACAGAGATGCGCATATGCTTCCAAAAGTAATGAGTGGGTTCGATTGACACACAA 1006
Qy 319 LysSerPheSerValLysAlaGlnTrpLeuLysGlnAsnAsnPheGlyGlyAlaMetIle 338
Db 1007 GAAAGTTATGAATGCAAGGTGAGTCTTAAAGAAAGTGGTTCGTTGGTGGCATGGTT 1066
Qy 339 TrpAlaIleAspLeuAspPheThrGlySerPheCysAspGlnGlyLysPheProLeu 358
Db 1067 TGGGCCATTCATTTGGATGATTTTCGAGGGGAGATTTTGCATCAAGAGAGATATCGTTG 1126
Qy 359 ThrSerThrLeuAsnLysAlaLeuGlyIleSerThrGluGlyCys----- 373
Db 1127 ATCAACACCTGAAGTCTTGTGGAGGTTCACCTGTAACCTGCGCTGGAGAAATATGT 1186
Qy 374 ----- 1199
Db 1187 GGTGGCATTTCTATACCACTCTCCAAGTCAACACCAACCAACCAACCAACAGCC 1246
Qy 375 AlaProAsp-----ValProSerGluProValThrProProGlySerGlySerGly 392
Db 1247 AACCTGACTGCACTACACAGAGCCTCCAGTCACTCCCTCCACT----- 1291
Qy 393 GlyGlySerSerGlyGlySerGlyGlySerGlyPheCysAlaAspLysAlaAspGly 412
Db 1292 -----GTGCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1345
Qy 413 LeuTyrProValAlaAspAspArgAsnAlaPheTrpGlnCysIleAsnGlyIleThrTyr 432
Db 1346 CTTTCATGTAATCCACTAAATACAAACAAATTTTACATTTGTGCAATGAGCTACTAT 1405
Qy 433 GlnGlnHisCysGlnAlaGlyLeuValPheAspThrSerCysAsnCysCysAsnTrpPro 452
Db 1406 TCAATGAATGTGCAGATGGCTTAGTATTTCAGGCAAGTTTGCACCTGTTGTAACGGCCA 1465

RESULT 15
AY065557
LOCUS AY065557
DEFINITION Mus musculus putative secretory protein precursor (Ym2) mRNA,
complete cds.
ACCESSION AY065557
VERSION AY065557.1 GI:18086513
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus.
REFERENCE
AUTHORS Chang,N.-C.A.
TITLE Tissue-specific Expression of Ym2 Protein
JOURNAL Unpublished
REFERENCE
AUTHORS Chang,N.-C.A.
TITLE Direct Submission
JOURNAL Submitted (05-DEC-2001) Institute of Microbiology & Immunology,
National Yang-Ming University, #155 Sec. 2 Li-Nong St., Beitou,
Taipei 112, Taiwan

FEATURES
source
Location/Qualifiers
1..1209
/organism="Mus musculus"
/strain="BALB/c"
/db_xref="taxon:10090"
/clone="pGEMT-ez-ym2.s"
/issue_type="stomach"
1..1209
/gene="Ym2"
1..1209
/gene="Ym2"
/codon_start=1
/product="putative secretory protein precursor"
/protein_id="AAL5751.1"
/db_xref="GI:18086514"
/translation="MAKILIVTGLAILNLVOLGSSYQLMVCYVTWAKDRPTGSGFKPG
NIDPCLTHLYAFAGKMNNEITYLSEQDLRDYALNGLKDRNTEKTLKLLAIGWKFG
PAPSSVMSTPQNRQTFIKSVIRFLQRYDFGLNDLMOYPGRSPGPKDHLFSLVQ
EMKAFEBESTLNHPRILLSTGAGFDIVKSGYKIPELSQSLDIYQVHTYDLDPK
NGYTGNSPLKSPYDIKGSADLNVDISIITWKDHGAASEKLIYGFAYGHTFILSDP
SKNGIDPVSAGPPGKVTNEQGLLAYFEICTFLNEGATEIFDQTOEVPYAVLGNWV
GYDNVRSFKLKAQWLKONNLGAVVWPLDMDDFSSGFCHQGRFPLTTTLKRDNLNVHSA
SCKASVYRGEL"

BASE COUNT 326 a 274 c 282 g 327 t

ORIGIN
Alignment Scores:
Pred. No.: 6,11e-94 Length: 1209
Score: 1368.00 Matches: 244
Percent Similarity: 81.33% Conservative: 61
Best Local Similarity: 65.07% Mismatches: 70
Query Match: 54.87% Indels: 0
DB: 10 Gaps: 0

US-10-004-219B-9 (1-452) x AY065557 (1-1209)

Qy 1 TyrAsnLeuIleCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPhe 20
Db 64 TACCAGCTGATGTGTCTACTATACAGCTGGCTAAGGAGGAGGAGGAGGAGTTTC 123
Qy 21 LysProAspAspIleAsnProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet 40
Db 124 AAACCTGTAATATATGACCCCTGCTGTACTACCTGATCTATGCTTGTGGGATG 183
Qy 41 GlnAsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrLysAlaPheAsn 60
Db 184 AAGATATATGATGATCCTTACTTAAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAG 243
Qy 61 AspLeuLysAsnArgAsnSerLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPhe 80
Db 244 GGTCTGAAAGACAGGACACACTGAGCTAAATAAACTCTCTGCGCATTTGGAGGATGGAAGTTT 303
Qy 81 GlyThrAlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheIleThr 100
Db 304 GGACCTGCCCCGTTCAGTTCCTACTCTCTCAGAACCCGTCAGACATTCATTAAAG 363
Qy 101 SerValIleLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGlyTyr 120
Db 364 TCAGTTATCAGATTCCTTCGTCATATATACTTTGATGGCTCAACCTGAGCTGGCAGTAC 423
Qy 121 ProGlySerArgGlySerProProGlnAspLysHisLeupheThrValLeuValLysGlu 140
Db 424 CTTGGGCTTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 483
Qy 141 MetArgGluAlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThr 160
Db 484 ATGCGTAAAGCTTTTGAGGAGAGATCCACTTTGAAACCACTTCCAAAGGCTGCTACTACT 543
Qy 161 AlaAlaValAlaGlyGlyIleSerAsnIleGlnAlaGlyTyrGluIleProGluLeuSer 180
Db 544 TCCACAGGAGCTGGATTCATTGACGTAATCAAGCTGGGTACAGATCCCTGNACTGCT 603
Qy 181 LysTyrLeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr 200

[illegible]

Search completed: June 29, 2003, 23:53:19
Job time : 2968 secs

THIS PAGE BLANK (USPTO)

XX 14-NOV-2000; 2000WO-JP08015.
 XX 15-NOV-1999; 99JP-0324467.
 XX (TAKE) TAKEDA CHEM IND LTD.
 XX Nakanishi A, Morita S;
 XX WPI; 2001-397791/42.
 XX New proteins, peptides and DNA for treatment of bronchial asthma.
 XX PT Chronic occlusive lung disease and infectious disease
 XX PS Claim 5; Page 100; 114pp; Japanese.
 XX The present invention provides the sequence of a protein which can be
 CC used in the treatment and prevention of infectious diseases. Inhibitors
 CC of the protein can be used to treat bronchial asthma and chronic
 CC occlusive pulmonary disease. The present sequence is an oligonucleotide
 CC described in the exemplification of the invention.
 XX SQ Sequence 1368 BP; 324 A; 398 C; 349 G; 297 T; 0 other;

Alignment Scores:
 Pred. No.: 3-22e-178 Length: 1368
 Score: 2089.50 Matches: 369
 Percent Similarity: 90.09% Conservative: 40
 Best Local Similarity: 81.28% Mismatches: 42
 Query Match: 83.81% Indels: 3
 DB: 22 Gaps: 1

US-10-004-219b-9 (1-452) x AAH42013 (1-1368)

QY 1 TyrAsnLeuLeuCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPhe 20
 DB 1 TACCAGCTGACATGCTACTTACCAACTGGGCCAGTACCGGCCAGGCTGGGGCGCTTC 60
 QY 21 LysProAspAspLeuAsnProCysLeuCysThrHisLeuLeuTyrAlaPheAlaGlyMet 40
 DB 61 ATGCCTGGAACAACATGACCCCTGCTGTACCCACTGATCCTGATCAGCCCTTGTCTGGGAGG 120
 QY 41 GlnAsnAsnGluLeuThrThrLeuGluTrpAsnAspValThrLeuTyrLysAlaPheAsn 60
 DB 121 CAGAACACAGAGATCACCACTGATGGAATGATGATGATGATGATGATGATGATGATGAT 180
 QY 61 AspLeuLysAsnArgAsnSerLysLeuLysThrLeuLeuAlaLeuGlyGlyTrpAsnPhe 80
 DB 181 GGCCCTGAAAAATAAGAACAGCCAGCTGAAAACTCTCTGCCCCATTGGAGGCTGGAACTTC 240
 QY 81 GlyThrAlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheLeuThr 100
 DB 241 GGAGCTGCCCTTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
 QY 101 SerValIleLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyr 120
 DB 301 TCAGTCATCAATTCCTGGCCAGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
 QY 121 ProGlySerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValLysGlu 140
 DB 361 CTGGCTCTCTGGGAGCCCTCTCAGGACCAAGCATCTCTTCACTGCTGCTGCTGCTGCTGCT 420
 QY 141 MetArgGluAlaPheGluGlnGluAlaLeuGluSerAsnArgProArgLeuMetValThr 160
 DB 421 ATGCGTGAAGCTTTTGAAGAGAGGCCAAGCATCAACAGCCAGGCTGATGGTGCAT 480
 QY 161 AlaAlaValAlaGlyGlyIleSerAsnIleGlnAlaGlyTyrGluIleProGluLeuSer 180
 DB 481 GCTGCACTAGCTGCTGGCATCTCCATATCCAGTCTGGCTATGAGATCCCCCACTGTCA 540
 QY 181 LysTyrLeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr 200
 DB 541 CAGTACCTGGACTTACATCCATGCTCATGACCTAGGACCTCCTGCTGCTGCTGCTGCTGCTGCT 600

QY 201 ThrGlyGluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeu 220
 DB 601 ACTGGAGAGACAGCCCTCTCAAAATACCGACTGACACCGGAGCAGACCCCTACCTC 660
 QY 221 AsnValAspTyrValMetAsnTyrTrpLysAsnAsnGlyAlaProAlaGluLysLeuIle 240
 DB 661 AATGTGGATTATGTATGAACCTACTTGAAGGACAATGGAGCACCAGCTGAGAAGCTCATC 720
 QY 241 ValGlyPheProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIle 260
 DB 721 GTTGAATTCCTTACCTATGGACACAACTTCACTCTGAGCAACCCCTCCAACTGGAAT 780
 QY 261 GlyAlaProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrp 280
 DB 781 GTGTGCCCCCACTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
 QY 281 AlaTyrTyrGluIleCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSer 300
 DB 841 GCTTACTAGAGATCTGTACCTTCTGAAAAATGAGGCACTCAGGGATGGATGCCCT 900
 QY 301 GlnGluValProTyrAlaTyrLysAlaAsnGluTrpLeuGlyTyrAspAsnIleLysSer 320
 DB 901 CAGGAAGTGCCTTATGCCTATCAGGGCAATGTGTGGTGTGCTGCTGCTGCTGCTGCTGCT 960
 QY 321 PheSerValLysAlaGlnTrpLeuLysGlnAsnAsnPheGlyGlyValaMetIleTrpAla 340
 DB 961 TTCGATATTAAAGCTCAATGCTTAAGCACACAATAATTTGGAGGCCCATGTGTGGGCC 1020
 QY 341 IleAspLeuAspAspPheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSer 360
 DB 1021 ATTGATCTGGATGACTTCACTGGCACTTCTGCAACACAGGCAAGTTTCCCTAATCTCC 1080
 QY 361 ThrLeuAsnLysAlaLeuGlyIleSerThrGluGlyCysThrAlaProAspValProSer 380
 DB 1081 ACCCTGAAAGAGCCCTCGCCCTGCAGAGTCAAGTTGCAAGCTCCAGCTCAGGCCCAT 1140
 QY 381 GluProValThrThrProPro-----GlySerGlySerGlyGlyGlySerGly 397
 DB 1141 GAGCCAAATAACTGCTGCTCCAGTGGCAGCGGAGACGGAGCGGAGTAGTAGCTCTGGA 1200
 QY 398 GlySerSerGlyGlySerGlyPheCysAlaAspLysAlaAspGlyLeuTyrProValAla 417
 DB 1201 GGCAGCTCGGAGGAGTGGATTCTGTGTGTCAGAGCAACAGGCTCTTACCCCTGGCA 1260
 QY 418 AspAspArgAsnAlaPheTrpGlnCysIleAsnGlyIleThrTyrGlnGlnHisCysGln 437
 DB 1261 AATAACAGAAAAATGCTTCTGGCACTGCGTGAATGGAGTCACTACACAGCAACTGCCAG 1320
 QY 438 AlaGlyLeuValPheAspThrSerCysAsnCysCysAsnTrp 451
 DB 1321 GCCGGCTTGTCTTCGACACACAGCTGTGATGCTGCAACTGG 1362

RESULT 2
 ID AAH42025
 AC AAH42025;
 DT 24-AUG-2001 (first entry)
 DE Disease treatment related oligonucleotide SEQ ID NO: 16.
 KW Disease treatment; infection; chronic occlusive pulmonary disease;
 KW bronchial asthma; ds.
 OS Homo sapiens.
 PN WO200136633-A1.
 XX 25-MAY-2001.
 XX 14-NOV-2000; 2000WO-JP08015.

XX 11-SEP-2000; 2000US-0659671.
 XX (HYSE-) HYSEQ INC.
 XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
 PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;
 XX WPI: 2002-292408/33.
 DR P-PSDB; ABB97580.
 XX
 PT An isolated polynucleotide for treating diseases associated with its
 PT encoded polypeptide such as cancer and multiple sclerosis -
 PS Claim 1; SEQ ID NO 404; 509pp; English.
 XX
 CC The present invention provides the protein and coding sequences of 444
 CC novel human proteins. These were isolated from expressed sequences tags
 CC (ESTs). They can be used to stimulate cell growth, to regulate
 CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
 CC e.g. in burn treatment, to regulate the immune system e.g. to treat
 CC multiple sclerosis, to regulate activin or inhibin e.g. to treat
 CC infertility, to regulate haemostasis or thrombolysis e.g. to treat
 CC stroke and cancer, to screen for drugs, to treat inflammatory conditions
 CC e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.
 CC Parkinson's disease. The present sequence is a coding sequence of the
 CC invention.
 XX
 SQ Sequence 1369 BP; 327 A; 393 C; 341 G; 308 T; 0 other;

Alignment Scores:
 Pred. No.: 2,88e-150 Length: 1369
 Score: 1778.50 Matches: 315
 Percent Similarity: 89.14% Conservative: 38
 Best Local Similarity: 79.55% Mismatches: 40
 Query Match: 71.34% Indels: 3
 DB: 24 Gaps: 1

US-10-004-219B-9 (1-452) x ABB95993 (1-1369)

QY	59	PheAsnAspLeuLysPheAsnArgSerLysLeuLysThrLeuLeuAlaIleGlyGlyTyr	78
DB	67	TATATGAAATTTGCACCTCGGACACGCCAGCTGAAACCTCTCTGCGCATTTGGAGCTGG	126
QY	79	AsnPheGlyThrAlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPhe	98
DB	127	AACTTCGGAGCTGCCCCCTTCACTGCCAGTGGTTTCTACTCTCGAGAACCGCCAGCTTTC	186
QY	99	IleThrSerValIleLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTyr	118
DB	187	ATCACCTCAGTCATCAATTTCTGCGCCAGTATGAGTTTGACGGGCTGGACTTTGACTGG	246
QY	119	GluTyrProGlySerArgGlySerProGlnAspLysHisLeuPheThrValLeuVal	138
DB	247	GAGTACCTCTGCTCTGCGGAGCCCTCTCAGGACAAGCATCTCTTCACTGCTCTGCTGGTG	306
QY	139	LysGluMetArgGluAlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMet	158
DB	307	CAGGAATTCGGTGAAGCTTTTGAGCAGGAGGCGAACGATCAACAGCCAGGCTGATG	366
QY	159	ValThrAlaAlaValAlaGlyGlyIleSerAsnIleGlnAlaGlyTyrGluIleProGlu	178
DB	367	GTCACTGCTGAGTAGCTGCTGGCATCTCCAATATCCAGTCTGGCTATGAGATCCCCAA	426
QY	179	LeuSerLysTyrLeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTrpGlu	198
DB	427	CTGTACAGTACTGAGTACATCCATGTCATGACCTACGACCTCCATGGCTCTCTGGGAG	486
QY	199	GlyTyrThrGlyGluAsnSerProLeuTyrLysTyrProThrGlySerAsnAla	218
DB	487	GACTACACTGGAGAGNACAGCCCTCTTCAAAATACCCGACTGACACCGGAGCAAGCC	546
QY	219	TyrLeuAsnValAspTyrValMetAsnTyrTrpLysAsnAsnGlyAlaProAlaGluLys	238

DB	547	TACCTCAATGTGGATTATGTATGAACACTACTGGAAGGACAATGGAGCACCACTGAGAAG	606
QY	239	LeuIleValGlyPheProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsn	258
DB	607	CTCATCTGTGGATTCCTTACCTATGAGACACAACTTCATCTGAGCAACCCCTCCCAACT	666
QY	259	GlyIleGlyAlaProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGly	278
DB	667	GGAAATTTGGTGGCCCACTCTGGTCTGTCTGGCCCTATGCCAAGAGTCTGGG	726
QY	279	PheThrAlaTyrTyrGluIleCysThrPheLeuArgSerGlyAlaThrGluValTrpAsp	298
DB	727	ATCTGGGCTTACTACGAGATCTGTACCTTCTCTGAAAAATGGAGCCACTCAGGATGGAT	786
QY	299	AlaSerGlnGluValProTyrAlaTyrLysAlaAsnGluTrpLeuGlyTyrAspAsnIle	318
DB	787	GGCCCTCAGGAAGTGCCTTATGCTTATCAGGCAATGTGGGTGGTGGTATGACAACTC	846
QY	319	LysSerPheSerValLysAlaGlnTrpLeuLysGlnAsnAsnPheGlyGlyAlaMetIle	338
DB	847	AGAGCTTCGATATTAAAGCTCAATGGCTTAAGCACACAAATTTGGAGCGCCATGGTC	906
QY	339	TrpAlaIleAspLeuAspAspPheThrGlySerPheCysAspGlnGlyLysPheProLeu	358
DB	907	TGGGCAATGATCTGGATGACTTCACTGGCACTTCTGCAACAGGCAAGTTTCCCTTA	966
QY	359	ThrSerThrLeuAsnLysAlaLeuGlyIleSerThrGluGlyCysThrAlaProAspVal	378
DB	967	ATCTCCACCTCGAAGAGGCGCTCGGCTGCGAGTGCAGGTTGCACGCTCCAGCTCAG	1026
QY	379	ProSerGluProValThrThrProPro-----GlySerGlySerGlyGlySer	395
DB	1027	CCCATTTAGCCAAATTAACCTGCTCCAGTGGCAGCGGAGCGGAGCGGAGTAGCAGC	1086
QY	396	SerGlyGlySerSerGlyGlySerGlyPheCysAlaAspLysAlaAspGlyLeuTyrPro	415
DB	1087	TCTGGAGGCGAGCTCGGAGGCGAGTGGATTCTGTGCTGTCAGAGCCACGCGCTCTACCC	1146
QY	416	ValAlaAspAspArgAsnAlaPheThrGlnCysIleAsnGlyIleThrTyrGlnGlnHis	435
DB	1147	GTGGCAAAATAACAGAAATGCTTCTGGCACTGCGTGAATGGAGTCACTGACGACAGAAC	1206
QY	436	CysGlnAlaGlyLeuValPheAspThrSerCysAsnCysCysAsnTyr	451
DB	1207	TGCCAGCGCGGCTGTCTTTCGACACCACTGTGATTGCTGCAACTGG	1254
RESULT 4			
ID	AAH42023		
XX	AAH42023	standard; DNA; 1469 BP.	
AC	AAH42023;		
XX	24-AUG-2001	(first entry)	
DE	Disease treatment related oligonucleotide SEQ ID NO: 14.		
KW	Disease treatment; infection; chronic occlusive pulmonary disease;		
XX	bronchial asthma; ds.		
OS	Mus sp.		
PN	WO200136633-A1.		
PD	25-MAY-2001.		
XX	14-NOV-2000;	2000WO-JP08015.	
PR	15-NOV-1999;	95JP-0324467.	
PA	(TAKE) TAKEDA CHEM IND LTD.		
PI	Nakanishi A, Morita S;		

XX WPI; 2001-397791/42.
 XX New proteins, peptides and DNA for treatment of bronchial asthma,
 PT chronic occlusive lung disease and infectious disease
 XX Disclosure; Page 106; 114pp; Japanese.
 XX The present invention provides the sequence of a protein which can be
 CC used in the treatment and prevention of infectious diseases. Inhibitors
 CC of the protein can be used to treat bronchial asthma and chronic
 CC occlusive pulmonary disease. The present sequence is an oligonucleotide
 CC described in the exemplification of the invention.
 XX
 XX Sequence 1469 BP; 406 A; 331 C; 332 G; 400 T; 0 other;
 SQ

Alignment Scores:
 Pred. No.: 4,86e-115 Length: 1469
 Score: 1387.00 Matches: 249
 Percent Similarity: 80.8% Conservative: 55
 Best Local Similarity: 66.22% Mismatches: 72
 Query Match: 55.64% Indels: 0
 DB: 22 Gaps: 0

US-10-004-219B-9 (1-452) x AAH42023 (1-1469)

QY 1 TyrAsnLeuLeuCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPhe 20
 DB 72 TACCAGCTGATGTCTACTATACCTAGTGGCTAGGACAGGCCAATAGAGGGAGTTTC 131

QY 21 LysProAspAspIleAsnProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet 40
 DB 132 AAACCTGGTAATATTGACCCCTGCTGTACTACCTGATCTATGCTTCTGGAATG 191

QY 41 GlnAsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrLysAlaPheAsn 60
 DB 192 CAGAAATATGAGATCATTACACATGATGACGAGCAAGCTTGGCTGATGATGAAGCTGAAT 251

QY 61 AspLeuLysAsnArgAsnSerLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPhe 80
 DB 252 GGTCTGAAAGACAGACAGACTAGCTGATTAACCTCTCTGGCCATTGGAGATGGAGTTT 311

QY 81 GlyThrAlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheIleThr 100
 DB 312 GGACCTGCCCTCTCAGTCCCATGCTCTACTCTCTCAGAACCGTCAGATATTCATTGAC 371

QY 101 SerValIleLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyr 120
 DB 372 TCAGTTATCAGATTCCTTCGTCATATATACCTTTGATGGCTCAACCTGGACTGGCAGTAC 431

QY 121 ProGlySerArgGlySerProGlnAspLysHisLeuPheThrValLeuValLysGlu 140
 DB 432 CTGGGTCTCGAGGAGCCCTCTAGGACAAACATCTCTCAGTGTCTGGTGAAGGAA 491

QY 141 MetArgGluAlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThr 160
 DB 492 ATGCGTAAAGCTTTTGAGGAGAACTGTGGAGAAAGACATTCGAAGCTGCTACTACT 551

QY 161 AlaAlaValAlaGlyGlyIleSerAsnIleGlnAlaGlyTyrGluIleProGluLeuSer 180
 DB 552 TCCACAGGAGCAGGAATCATTCACGTAATCAAGTCTGGGTACAAAGATCCCTGAACCTGCT 611

QY 181 LysTyrLeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr 200
 DB 612 CAGTCTCTTGACTATATCAGGTGATGATATGATATGATATGATATGATATGATATGATAT 671

QY 201 ThrGlyGluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeu 220
 DB 672 ACTGGAGAAAATAGTCCCTCTATAAATCTCCATGATGATGATGATGATGATGATGAT 731

QY 221 AsnValAspTyrValMetAsnTyrTrpLysAsnAsnGlyAlaProAlaGluLysLeuIle 240
 DB 732 AATGTGGATTCAATCTTCTACTGGAAGGACCATGGAGCTTCTGGAAGCTCAT 791

QY 241 ValGlyPheProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIle 260
 DB 792 GTGGATTTCCAGCATATGGGCATACCTTTATCTCTGAGTGACCTTCTAAGACTGGAATT 851

QY 261 GlyAlaProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrp 280
 DB 852 GGTGCCCTTACAATTTAGTACTGGCCACCCAGGAAAGTACACAGATGAATCAGGACTCCTG 911

QY 281 AlaTyrTyrGluIleCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSer 300
 DB 912 GCTTACTATAGGTTTGTACATTTCTGAATGAAGGAGCCACTGAGGTCTGGATGCCCCC 971

QY 301 GlnGluValProTyrAlaTyrLysAlaAsnGluTrpLeuGlyTyrAspAsnIleLysSer 320
 DB 972 CAGGAAGTACCTATGCTATCAGGGTAAATGAGTGGTGGTTGTTATGACAATGTCCAGGAC 1031

QY 321 PheSerValLysAlaGlnTrpLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTrpAla 340
 DB 1032 TTCAGTTGAAAGGCTCAGTGCTCAAGGACAAACAATTTAGGAGGTGCGGTGCTGGCCC 1091

QY 341 IleAspLeuAspAspPheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSer 360
 DB 1092 CTGGACATGGATGACTTCAGTGGTTCTTCTGTCTACCAGAGACATTCCTCTGACATCT 1151

QY 361 ThrLeuAsnLysAlaLeuGlyIleSerThrGluGlyCysThrAlaPro 376
 DB 1152 ACTTTAAGGGAGATCTCAATATACACAGTGCAGTGCAGGGGCCCT 1199

RESULT 5
 AAIV10435
 ID AAV10435 standard; cDNA; 1636 BP.
 XX
 AC AAV10435;
 XX
 DT 15-JUN-1998 (first entry)
 XX
 DE Human chitinase clone MO-218 cDNA.
 XX
 KW Chitinase; human; fungal infection; immunogen; diagnosis; treatment;
 KW Gaucher's disease; transgenic; detection; hybridisation; antifungal;
 KW Rheumatoid arthritis; overexpression; extracellular matrix; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 2..1402
 FT /*tag= a
 FT sig_peptide 2..64
 FT /*tag= b
 FT mat_peptide 65..1399
 FT /*tag= c
 FT /product= chitinase
 FT /note= "from clone MO-218"
 XX
 XX W09747752-A1.
 XX
 XX PN
 XX
 PD 18-DEC-1997.
 XX
 XX 16-JUN-1997; 97WO-US10460.
 XX
 XX 14-JUN-1996; 96US-0663618.
 XX
 XX (ICOS-) ICOS CORP.
 XX
 XX Gray PW;
 XX
 XX WPI; 1998-052316/05.
 XX P-PSDB; AA40259.
 XX
 XX Nucleic acids encoding human chitinase - useful as antifungal
 XX agents, especially in combination with other antifungals
 XX

/product= Mature_protein

FT XX W09946390-A1.
 PN XX
 XX
 PD 16-SEP-1999.
 PF 12-MAR-1999; 99WO-US05343.
 XX
 PR 12-MAR-1998; 98US-0039198.
 XX
 PA (ICOS-) ICOS CORP.
 XX
 PI Gray PW, Tjoelker LW;
 XX
 DR WPI; 1999-551417/46.
 DR P-PSDB; AAY42425.
 XX
 FT Novel chitin-binding fragments of human chitinase used to treat fungal
 PT infections in animals -
 XX
 PS Example 1; Page 55-57; 83pp; English.
 CC This is the nucleotide sequence of an allelic form of the human
 CC chitinase enzyme, which is capable of degrading Chitin (a linear
 CC homo polymer of beta-1,4-linked N-acetylglucosamine residues).
 CC Chitinase fragments can be used to screen for proteins or other
 CC molecules that specifically bind to the chitin-binding domain of human
 CC chitinase or that modulate its activity. These compounds are useful for
 CC immunization, as well as for purifying chitinase, as well as for
 CC detection and quantification of chitinase. Polynucleotide fragments of
 CC the invention are useful as a source of probes and primers, and to
 CC express the proteins recombinantly. The chitinase fragments, when
 CC conjugated to antifungal compounds, are used to treat animals, such
 CC especially humans, infected with chitin-containing parasites such as
 CC fungi. Fungal infection treated include candidiasis, aspergillosis,
 CC coccidioidomycosis, blastomycosis, paracoccidioidomycosis,
 CC mucormycosis, histoplasmosis, cryptococcosis, chromoblastomycosis,
 CC sporotrichosis, and dermatophycoses.
 CC Chitin can be degraded by the enzyme chitinase. Use of whole chitinase
 CC protein for treating infections, especially fungal infections, is
 CC problematic. In view of the increasing incidents of life-threatening
 CC fungal infection in e.g. immunocompromised individuals, there exists a
 CC need for identifying new compounds for treating fungal infection. The
 CC chitin-binding fragments of the present invention provide this need.
 XX
 SQ Sequence 1636 BP; 361 A; 491 C; 440 G; 344 T; 0 other;

Alignment Scores:

Pred. No.:	2,71e-107	Length:	1636
Score:	1301.50	Matches:	235
Percent Similarity:	67.11%	Conservative:	71
Best Local Similarity:	51.54%	Mismatches:	129
Query Match:	52.21%	Indels:	21
DB:	20	Gaps:	3

US-10-004-219B-9 (1-452) x AAZ21847 (1-1636)

QY 3 LeuileCysThrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPheLysPro 22
 DB 71 CTGGTCTGCTACTTACCAACTGGGCCCAAGTACAGACAGGGGGAGGCTCGCTTCCTGCC 130
 QY 23 AspAspIleAsnProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyMetGlnAsn 42
 DB 131 AAGGACTTGGAGCCCGCCAGCTTTGGACCCCACTCACTACGCTTCGCTGGCATGACCAAC 190
 QY 43 AsnGluileThrIleGlnTrpAsnAspValThrLeuTyrLysAlaPheAsnAspLeu 62
 DB 191 CACCACTGAGCACCACCTGAGTGGATGACGAGACTCTCTACCAAGAGTTCATGSCCTG 250
 QY 63 LysAsnArgAsnSerLysLeuIleThrLeuAlaIleGlyTyrTrpAsnPheGlyThr 82
 DB 251 AAGAAGATGAATCCCAAGCTGAAGACCTGTTAGCCATCGGAGGCTGGAAATTTCCGCACT 310

QY 83 AlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheIleThrSerVal 102
 DB 311 CAGAAGTTTACAGATATGTTAGCCACCGCCCAACACCTTGTCACTCGGCC 370
 QY 103 IleLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyrProGly 122
 DB 371 ATCAGGTTTCTGCGCAATACAGCTTTGACGCGCTTGACCTTGAGTGGAGTACCAGGA 430
 QY 123 SerArgGlySerProGlnAspLysHisLeuPheThrValLeuValLysGluMetArg 142
 DB 431 AGCCAGGGAGCCCTGCGGTAGACAGAGGCGCTTCAACCTGGTACAGACTTGGCC 490
 QY 143 GluAlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThrAlaAla 162
 DB 491 AATGCTTCCAGCAGGAAGCCAGACCTCAGGAAGGAACGCTTCTTCTGAGTGCAGCG 550
 QY 163 ValAlaGlyGlyLysSerAsnIleGlnAlaGlyTyrGluIleProGluLeuSerLysTyr 182
 DB 551 GTTCCAGCTGGGCGAGACCTATGTGGATGCTGGATACGAGGTGGACAAATCGCCAGAAC 610
 QY 183 LeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTyrGluGlyTyrThrGly 202
 DB 611 CTGGATTTGTCAACCTTATGCGCTTACGACTTCCATGCTCTTGGGAGAAGTCAAGGA 670
 QY 203 GluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeuAsnVal 222
 DB 671 CATAACACCCCTCTACAGAGCAAGAGAGTGTGCGAGCAGCCAGCCTCAACGTG 730
 QY 223 AspTyrValMetAsnTyrTrpLysAsnGlyAlaProAlaGluLysLeuIleValGly 242
 DB 731 GATGCTGTGTCAACAGTGTGCGAGAGGAGGAGCCCTGCGCAGCAAGTATCTTGGC 790
 QY 243 PheProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIleGlyAla 262
 DB 791 ATGCTTACTACGAGCGCTCTTCACTGGCTCTCATCAGACACCGAGTGGGGGCC 850
 QY 263 ProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrpAlaTyr 282
 DB 851 CCAGCCACAGGGTCTGGCAGCTCCAGGCCCTTCCACCAAGAGGAGGAGTGTGGCTTAC 910
 QY 283 TyrGluIleCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSerGlnGlu 302
 DB 911 TATGAAGTCTGCTCTCTGG-----AAGGGGGGCCCAACCAACAGAGATCCAGGATCAGAAG 964
 QY 303 ValProTyrAlaTyrLysAlaAsnGluTrpLeuGlyTyrAspAsnIleLysSerPheSer 322
 DB 965 GTGCCCTACATCTTCCGGGCAACCCAGTGGGTGGGCTTTGATGATGGAGAGCTTCAAA 1024
 QY 323 ValLysAlaGlnTrpLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTrpAlaIleAsp 342
 DB 1025 ACCAAGTTCAGCTATCTGACAGCAAGAGGAGTGGGGGGGCCCATGCTTGGGCACTGGAC 1084
 QY 343 LeuAspAspPheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSerThrLeu 362
 DB 1085 TTAGATGACTTTCGCGGCTTCTCTCTGCAACCAAGGCGCCGATACCCCTCATCCAGACGCTA 1144
 QY 363 AsnLysAlaLeuGlyIleSerThrGluGlyCysThrAlaProAspVal----- 378
 DB 1145 CGGCAGGAAGTCTGAGCTTCCATCTTCCAGTCTTCCAGCACCCCGAGAGCTTGAAGTCCAAA 1204
 QY 379 -----ProSerGluProValThrThrProGlySerGlySerGlyGlyGlySer 395
 DB 1205 CCAGTTCAGCCCTCTGAACCT-----GAG 1228
 QY 396 SerGlyGlySerSerGlyGlySerGlyPheCysAlaAspLysAlaAspGlyLeuTyrPro 415
 DB 1229 CATGGCCCCAGCCCTGGCAAGACACAGCTTCTGCCAGGCGCAAGCTCATGGGCTTATCCC 1288
 QY 416 ValAlaAspAspArgAsnAlaPheTrpGlnCysIleAsnGlyIleThrTyrGlnGlnHis 435
 DB 1289 AATCTCGGGAACGGTCCAGCTTCTACAGCTGTGCGGGGGCGGCTGTTCAGCAAGAAC 1348
 QY 436 CysGlnAlaGlyLeuValPheAspThrSerCysAsnCysCysAsnTrp 451

Db 1349 TGCCGACAGGCGTGTGTTACAGCACTCTCTGCAAAATGTCACCTGG 1396

RESULT 7

AD03759

ID AAD03759 standard; cDNA; 1636 BP.

AC AAD03759;

DT 19-JUN-2001 (first entry)

XX Human chitinase cDNA from clone pMO-218.

DE

KW Human; antifungal; chitinase; immunoglobulin; Ig; therapy;

KW fungal infection; candidiasis; aspergillosis; coccidioidomycosis;

KW blastomycosis; paracoccidioidomycosis; histoplasmosis; cryptococcosis;

KW chromoblastomycosis; sporotrichosis; mucormycosis; dermatophytoses;

KW clone pMO-218; ss.

XX

OS Homo sapiens.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Best Local Similarity: 51.54%

Query Match: 52.21%

DB: 22

Mismatches: 129

Indels: 21

Gaps: 3

US-10-004-219B-9 (1-452) x AAD03759 (1-1636)

QY 3 LeuileCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPheLysPro 22
 DB 71 CTGGTCTGTACTTCCACCACTGGGCCCGAGTACAGACAGGGGGAGGCTCGCTTCTGCCC 130
 QY 23 AspAspIleAsnProCysLeuThrsHisLeuileTyrAlaPheAlaGlnMetGlnAen 42
 DB 131 AAGGACTTGGACCCCGACCTTTGGACCCCACTCATCTACGCTTCGCTGGCATGCCAAC 190
 QY 43 AenGluileThrThrIleGluTrpAsnAspValThrLeuTyrLysAlaPheAsnAspLeu 62
 DB 191 CACCACTGAGCACCACTGAGTGGATGACGAGACTCTCTACCAAGAGTTCAATGGCCTG 250
 QY 63 LysAsnArgAsnSerLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPheGlyThr 82
 DB 251 AAGAAGATGAATCCCAAGCTGAAGACCTGTTAGCATCGGAGGCTGGAATTTCCGCACT 310
 QY 83 AlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheIleThrSerVal 102
 DB 311 CAGAAATTCACAGATATGTTAGCCACGCGCAACACCTCTAGACTTTGTCAACTCGGCC 370
 QY 103 IleLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyrProGly 122
 DB 371 ATCAGTTTCTGGCAATATACAGCTTTGACGCGCTTGACCTTGAGTGGAGTACCAGGA 430
 QY 123 SerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValLysGluMetArg 142
 DB 431 AGCCAGGGGAGCCCTGCGGTAGACAAAGAGGCGCTTCACAACTCGGTACAGACTTGGCC 490
 QY 143 GluAlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThrAlaAla 162
 DB 491 AATGCTTTCAGCAGGAAGCCAGACCTCAGAGGAAGAACGCTTCTTCTAGTGCAGCG 550
 QY 163 ValAlaGlyGlyIleSerAsnIleGlnAlaGlyTyrGluileProGluLeuSerLysTyr 182
 DB 551 GTTCCAGCTGGGCAGACCTATGTGATGCTGGATAGAGGTGGACAAATCGCCGAGAAC 610
 QY 183 LeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThrGly 202
 DB 611 CTGGATTTGTCACTTATGCGCTTACGACTTCCATGCTCTTGGGAGAAGGTCACGGGA 670
 QY 203 GluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeuAenVal 222
 DB 671 CATAACACCCCTCTACAGAGGCAAGAGAGTGTGTGAGCAGCAGCGCTCAACCTG 730
 QY 223 AspTyrValMetAsnTyrTrpLysAsnAsnGlyAlaProAlaGluLysLeuileValGly 242
 DB 731 GATGCTGTGTGCAACAGTGGCTGCAGAGGGGACCCCTGCCAGCAAGCTGATCTTGGC 790
 QY 243 PheProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIleGlyAla 262
 DB 791 ATGCTACTACTGAGGAGCTCTTCTACACTGGCTCTCTCATCAGACACCAAGAGTGGGGCC 850
 QY 263 ProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrpAlaTyr 282
 DB 851 CCAGCCACAGAGTCTGGGCACTCCAGGCGCTTCCAAAGGAGGAGGATGTCGCTGCTAC 910
 QY 283 TyrGluileCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSerGlnGlu 302
 DB 911 TATGAGTCTGCTCTCTGG-----AAGGGGGCCCAACACAGAGATCCAGGATCAGAAG 964
 QY 303 ValProTyrAlaTyrLysAlaAsnGluTrpLeuGlyTyrAspAsnIleLysSerPheSer 322
 DB 965 GTGCCCTACATCTCCGGGCAACACAGTGGGTGGGCTTTGATGATGTGGAGAGCTTCAAA 1024
 QY 323 ValLysAlaGlnTrpLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTrpAlaIleAsp 342
 DB 1025 ACCAAGGTTCAGCTATCTGAAGCAGAGGAGTGGGGGGGGCCATGCTCTGGCACTTGAC 1084

Novel chitinase immunoglobulin fusion product, useful for treating fungal infections and reducing the amount of a non-chitinase antifungal agent needed for the treatment

Claim 2; Page 29-31; 39pp; English.

The present invention relates to a chitinase immunoglobulin (Ig) fusion product, comprising a human chitinase fused to at least a portion of an immunoglobulin chain. The fusion product is useful for treating fungal infections (mycoses) such as candidiasis, aspergillosis, blastomycosis, coccidioidomycosis, paracoccidioidomycosis, histoplasmosis, mucormycosis, cryptococcosis, chromoblastomycosis, sporotrichosis and dermatophytoses. The fusion protein is useful for reducing the amount of non-chitinase antifungal agent needed to exert an antifungal activity. The fusion protein is also useful for preparing a medicament for the prophylactic or therapeutic treatment of fungal infections. Chitinase immunoglobulin or fusion product has unexpectedly improved serum half-life and formulation properties. The present sequence is human chitinase cDNA from clone pMO-218. Chitinase enzyme degrades chitin which is a homopolymer of beta-(1,4)-linked N-acetylglucosamine residues.

Sequence 1636 BP; 361 A; 491 C; 440 G; 344 T; 0 other;

Alignment Scores:

Pred. No.: 2,71e-107 Length: 1636

Score: 1301.50 Matches: 235

Percent Similarity: 67.11% Conservative: 71

QY 343 LeuAspAspPheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSerThrLeu 362
 DB 1085 TTAGATGACTTTGGCGGCTTCTCTGCAACAGGCGCGATACCCCTCATCCAGAGCTA 1144
 QY 363 AnLysAlaLeuGlyLysSerThrGluGlyCysThrAlaProAspVal 378
 DB 1145 CGGCGAACTTCACTTCTCCATCTTGCCTTCAGGCACCCAGAGCTTGAAGTTCACAAA 1204
 QY 379 -----ProSerGluProValThrThrProProGlySerGlySerGlyGlySer 395
 DB 1205 CCAGGTCAGCTCTGAACCT-----GAG 1228
 QY 396 SerGlyGlySerSerGlyGlySerGlyPheCysAlaAspLysAlaAspGlyLeuThrPro 415
 DB 1229 CATGGCCCGAGCCCTGGACACACAGCTTCTGCGAGGCAAAAGCTGATGGGCTCTATCC 1288
 QY 416 ValAlaAspAspArgAsnAlaPheTrpGlnCysIleAsnGlyIleThrTyrglnGlnHis 435
 DB 1289 ANTCTCGGGAAGCTCCAGCTTCTACAGCTGTGCAGCGGGCGGCTGTTCAGCAAGC 1348
 QY 436 CysGlnAlaGlyLeuValPheAspThrSerCysAsnCysCysAsnTrp 451
 DB 1349 TGCCCGACAGGCTGTGTTCAGCAACTCTGCAAAATGCTGCACCTGG 1396
 RESULT 8
 ABL57380
 ID ABL57380 standard; cDNA; 1636 BP.
 AC ABL57380;
 XX
 DT 12-AUG-2002 (first entry)
 XX
 DE Human chitinase cDNA clone MO-218.
 XX
 KW Chitinase; enzyme; human; fungicide; antifungal; infection;
 KW candidiasis; aspergillosis; coccidioidomycosis; blastomycosis;
 KW paracoccidioidomycosis; histoplasmosis; cryptococcosis;
 KW chromoblastomycosis; sporotrichosis; mucormycosis; dermatophytosis;
 KW Pneumocystis; gene; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 2..1402
 FT /tag= a
 FT /product= "Chitinase"
 FT sig_peptide 2..64
 FT /tag= b
 FT mat_peptide 65..1399
 FT /tag= c
 XX
 PN US6372212-B1.
 XX
 PD 16-APR-2002.
 XX
 PF 16-JUN-1997; 97US-0877599.
 XX
 PR 14-JUN-1996; 96US-0663618.
 XX
 PA (ICOS-) ICOS CORP.
 XX
 PI Gray PW;
 XX
 DR WPI; 2002-442449/47.
 DR P-PSDB; ABB76291.
 XX
 PT Co-administering chitinase to improve the effectiveness of fungicidal
 PT drugs e.g. amphotericin B or itraconazole, useful for treating fungal
 PT infections e.g. candidiasis, coccidioidomycosis and blastomycosis -
 XX Example 1; Column 23-28; 26pp; English.
 PS
 XX

CC The present sequence is that of cDNA clone MO-218 (ATCC 98077)
 CC encoding human chitinase (see ABB76291). The clone was isolated
 CC from a cDNA library prepared from peripheral blood monocyte-derived
 CC macrophages following sequence analysis. Also isolated was clone
 CC MO-13B (see ABL57381), which contains a single nucleotide difference
 CC in the coding region, changing the encoded amino acid at position 81
 CC of the mature protein from glycine to serine. Northern blots showed
 CC highest chitinase gene expression in lung and ovary tissues.
 CC Expression in lung is consistent with a protective role against
 CC pathogenic organisms that contain chitin. The invention provides
 CC human chitinase polynucleotides and polypeptides, and materials and
 CC methods for the recombinant production of human chitinase products,
 CC which are expected to be useful as products for treating fungal
 CC infections or for the development of such products. Human
 CC chitinase has a synergistic effect on the actions of other
 CC fungicides. It can be administered to improve the antifungal
 CC activity of a non-chitinase antifungal agent, especially
 CC amphotericin B or itraconazole, in the treatment of a fungal
 CC infection such as candidiasis, aspergillosis, coccidioidomycosis,
 CC blastomycosis, paracoccidioidomycosis, histoplasmosis,
 CC cryptococcosis, chromoblastomycosis, sporotrichosis, mucormycosis,
 CC dermatophytoses and Pneumocystis infections (all claimed). In
 CC particular, the fungal infection involves Candida, Aspergillus
 CC and/or Cryptococcus spp., whose growth is not effectively
 CC inhibited by contact with human chitinase alone.
 XX
 SQ Sequence 1636 BP; 361 A; 491 C; 440 G; 344 T; 0 other;
 Alignment Scores:
 Pred. No.: 2,71e-107 Length: 1636
 Score: 1301.50 Matches: 235
 Percent Similarity: 67.11% Conservative: 71
 Best Local Similarity: 51.54% Mismatches: 129
 Query Match: 52.21% Indels: 21
 DB: 24 Gaps: 3
 US-10-004-219b-9 (1-452) x ABL57380 (1-1636)
 QY 3 LeuIleCysTyThrPheThrAsnTrpAlaGlnTyArgProGlyLeuGlySerPheLysPro 22
 DB 71 CTGGTCTGCTACTTCCCACTGGGCCCGCCAGTACACAGGCGGAGGCTCGCTTCTCGCC 130
 QY 23 AspAspIleAsnProCysLeuCysThrHisLeuIleTyThrAlaPheAlaGlyMetGlnAsn 42
 DB 131 AAGGACTTGGACCCAGCTTTGACCCACCTCATCTAGCCCTTCGTCGCATGACCAAC 190
 QY 43 AsnGluIleThrIleGluTrpAsnAspValThrLeuTyLysAlaPheAsnAspLeu 62
 DB 191 CACCAGCTGAGCACCACTGAGTGGAAATGACGAGACTCTCTACCGAGGTTCATATGGCCTG 250
 QY 63 LysAsnArgAsnSerLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPheGlyThr 82
 DB 251 AAGAAGATGATCCCAAGCTGAAGACCTGTAGCATCGGAGGCTGGATTTGGCACT 310
 QY 83 AlaProPheThrMetValSerThrSerGlnAsnArgGlnThrPheIleThrSerVal 102
 DB 311 CAGAAGTTTCACAGATATGTTAGCCACGCGCCACACACCTTCAGACCTTTGTCACTCGGCC 370
 QY 103 IleLysPheLeuArgGlnTyThrGlyPheAspGlyLeuAspLeuAspTrpGluTyProGly 122
 DB 371 ATCAGGTTTCTCGCAAAATACAGCTTTGACGCCCTTGACCTTGAGCTGGGAGTACCCAGGA 430
 QY 123 SerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValLysGluMetArg 142
 DB 431 AGCCAGGGGAGGCCCTGCGGTAGCAAGGAGGCGCTTCACAAACCTTGATCAGGACTTTGGCC 490
 QY 143 GluAlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThrAlaAla 162
 DB 491 AATGCTTCCAGCAGAGAGCCAGACCTTCAGGAGGAGGAGGCTTCTTCTGAGTGCAGCG 550
 QY 163 ValAlaGlyGlyIleSerAsnIleGlnAlaGlyTyrgluIleProGluLeuSerLysTy 182
 DB 551 GTTCCAGCTGGGCGAGACCTATGTGGATGCTGGATACGAGGTGACAAATATCGCCAGAAC 610

XX	Key	Location/Qualifiers	
XX	CDS	13..1413	
FT		/*tag= a	
FT	sig_peptide	13..75	
FT		/*tag= b	
FT	mat_peptide	76..1410	
FT		/*tag= c	
XX	WO96040940-A2.		
XX	19-DEC-1996.		
XX	06-JUN-1996;	96WO-NL00225.	
XX	07-JUN-1995;	95US-0486839.	
XX	(UNAM) UNIV AMSTERDAM.		
XX	Aerts JWFG;		
XX	WPI; 1997-118698/11.		
XX	P-PSDB; AAW08584.		
XX	New human chitinase - used to treat or prevent infection by		
PT	chitin-contg. pathogens, in diagnosis and as additives to cosmetics,		
PT	foods, implants etc.		
XX	Claim 2; Page 39-40; 58pp; English.		
XX	A cDNA clone (AAT50833), designated chi.50, codes for a 50 kDa human		
CC	chitinase (AAW08584) that is stable to many proteases, active at pH 3-		
CC	8 and up to 50 deg, and stable in the circulation. Clones chi.50		
CC	and chi.39 (see also AAT50834) were isolated from a human macrophage		
CC	cDNA library using as probe a partial clone obtd. using primers		
CC	(see also AAT50835-36) based on a chitotriosidase purified from a		
CC	type 1 Gaucher disease patient. The 2 cDNA clones are the result		
CC	of alternative splicing of RNA. Chitinase nucleic acid can be used		
CC	for large-scale prodn. of recombinant human chitinases, or can be		
CC	incorporated into a gene therapy vector to treat or prevent		
CC	infection by chitin-contg. pathogens.		
XX	Sequence 1643 BP; 364 A; 490 C; 442 G; 347 T; 0 other;		
XX	Sequence 1643 BP; 364 A; 490 C; 442 G; 347 T; 0 other;		
XX	Alignment Scores:		
XX	Pred. No.:	2.73e-107	Length: 1643
XX	Score:	1301.50	Matches: 235
XX	Percent Similarity:	67.11%	Conservative: 71
XX	Best Local Similarity:	51.54%	Mismatches: 129
XX	Query Match:	52.21%	Indels: 21
XX	DB:	18	Gaps: 3
XX	US-10-004-219B-9 (1-452) x AAT50833 (1-1643)		
QY	3	LeulleCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPheLysPro	22
Db	82	CTGGTCTGTACTTTCACCAATGGGCCCAAGTACAGACAGGGGGAGGCTCGCTTCTGCGCC	141
QY	23	AspAspIleAsnProCysLeuCyThrHisLeulleTyrAlaPheAlaGlyMetGlnAsn	42
Db	142	AGGACTTGGACCCAGCCCTTTCACCCACCTCATCTACGCCCTTCGCTGGCATGCCAAC	201
QY	43	AenGluIleThrThrIleGluTrpAsnAspValThrLeuTyrIysAlaPheAsnAspLeu	62
Db	202	CACCAGCTGAGCACCACCTCAGTGGAAATGACGAGACTCTCTACCCAGGAGTTCAATGCGCTG	261
QY	63	LysAsnArgAsnSerIysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPheGlyThr	82
Db	262	AGAAGATGAATCCCAAGCTGAACCCCTTTAGCCATCGAGGCTGGAAATTCGGCACT	321
QY	83	AlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheIleThrSerVal	102
Db	322	CAGAAGTTTCAAGATATGTTAGCCAGGCCCAACACCGTCAGACCTTTGTCAACTCGGCC	381

```
103 IleLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyrProGly 122
104 .....
382 ATCAGGTTCTCGGCAATACAGCTTTCAGCGCTTACCTTGACTGGGAGTACCAGGA 441
105 .....
123 SerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValLysGluMetArg 142
106 .....
442 AGCCAGGGGAGCCCTGCCGTAGCAGAGGAGCGCTTCACAAACCTGGTACAGGACTTGGCC 501
107 .....
143 GluAlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThrAlaAla 162
108 .....
502 AATGCTCTCCAGAGAACCCAGACCTCAGGAAGAACGCTTCTTCGTAGTGAGCG 561
109 .....
163 ValAlaGlyGlyIleSerAsnIleGlnAlaGlyTyrGluIleProGluLeuSerLysTyr 182
110 .....
562 GTTCCAGCTGGGCAGACCTATGTGATGCTGTGATGAGTGGACAAATCCGCCAGAC 621
111 .....
183 LeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThrGly 202
112 .....
622 CTGGATTTTGTCAACCTTATGGCTACGACTTCCATGGCTCTTGGGAGAAGGTCACGGGA 681
113 .....
203 GluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeuAsnVal 222
114 .....
682 CATACAGCCCTCTACAAAGAGGCAAGAGAGTGGTGCAGACGACCTCACAGTG 741
115 .....
223 AspTyrValMetAsnTyrTrpLysAsnAsnGlyAlaProAlaGluLysLeuIleValGly 242
116 .....
742 GATGCTGCTGTCACAGTGGTGCAGAGGGGACCCCTGCCAGCAGCTGATCTTGGC 801
117 .....
243 PheProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIleGlyAla 262
118 .....
802 ATGCTTACCTACGAGCGCTCTTCACTGGCTCTCTCATCAGACACACAGAGTGGGGCC 861
119 .....
263 ProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrpAlaTyr 282
120 .....
862 CCAGCCAGAGGTGTGGCACTCAGCGCCCTTCCACCAAGAGAGGAGGATGCTGGCCTAC 921
121 .....
283 TyrGluIleCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSerGlnGlu 302
122 .....
922 TATGAGTGTGCTCTCGG-----AAGGGGGCCACCAACAGAGATCCAGGATCAGAG 975
123 .....
303 ValProTyrAlaTyrLysAlaGlnGluTrpLeuGlyTyrAspAsnIleLysSerPheSer 322
124 .....
976 GTGCCCTACATCTTCGGGACCAACAGTGGTGGGCTTTGATGATGGAGAGCTTCAAA 1035
125 .....
323 ValLysAlaGlnTrpLeuLysGlnAsnAsnPheGlyGlyValAlaMetIleTrpAlaIleAsp 342
126 .....
1036 ACCAAGGTGAGCTATCTGAAGCAGAGAGGAGCTGGGGGGGCGCATGCTTGGGCACTGGAC 1095
127 .....
343 LeuAspAspPheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSerThrLeu 362
128 .....
1096 TTAGATGACTTTCGGGCTTCTCTGCAACAGGGGCGGATACCCCTCATCAGAGCTA 1155
129 .....
363 AsnLysAlaLeuGlyIleSerThrGluGlyCysThrAlaProAspVal-----378
130 .....
1156 CGGCAAGAACTCAGCTTCCATCTTCCATCTCAGGCAACCCAGAGCTTGAAGTTCCAAA 1215
131 .....
379 -----ProSerLysProValThrThrProProGlySerGlySerGlyGlySer 395
132 .....
1216 CCAGGTGACCCCTCTGAACCT-----GAG 1239
133 .....
396 SerGlyGlySerSerGlyGlySerGlyPheCysAlaAspLysAlaAspGlyLeuTyrPro 415
134 .....
1240 CATGGCCCGAGCCCTGGACAGACAGCTTCTGCCAGGGCAAGAGCTGATGGGCTCTATCCC 1299
135 .....
416 ValAlaAspAspArgAsnAlaPheTrpGlnCysIleAsnGlyIleThrTyrGlnGlnHis 435
136 .....
1300 AATCTCGGGAAACGGTCCAGCTTCTACAGCTGTGACGGGGGGGGTGTTCACCAAGC 1359
137 .....
436 CysGlnAlaGlyLeuValPheAspThrSerCysAsnCysCysAsnTrp 451
138 .....
1360 TGCCCGACAGGCGCTGTGTTCAGCACTCTCTCAATGTGCACCTGG 1407
```

RESULT 10

```
AAV10436
ID AAV10436 standard; cDNA; 1656 BP.
XX
AC AAV10436;
XX
DT 15-JUN-1998 (first entry)
XX
DE Human chitinase clone MO-13B cDNA.
XX
KW Chitinase; human; fungal infection; immunogen; diagnosis; treatment;
KW Gaucher's disease; transgenic; detection; hybridisation; antifungal;
KW rheumatoid arthritis; overexpression; extracellular matrix; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 27..1427
FT /tag= a
FT sig_peptide 27..89
FT /tag= b
FT mat_peptide 90..1424
FT /tag= c
FT /product= chitinase
FT /note= "from clone MO-13B"
XX
XX W09747752-A1.
XX
PD 18-DEC-1997.
XX
PF 16-JUN-1997; 97WO-US10460.
XX
PR 14-JUN-1996; 96US-0663618.
XX
PA (ICOS-) ICOS CORP.
XX
PI Gray PW;
XX
DR WPI: 1998-052316/05.
DR P-PSDB; AAW40260.
XX
XX Nucleic acids encoding human chitinase - useful as antifungal
agents, especially in combination with other antifungals
XX
PS Claim 9; Page 42-44; 63pp; English.
XX
XX This sequence encodes a novel human chitinase isolated from clone MO-13B.
XX Chitinases are useful for treating or preventing fungal infection and
XX as immunogens for generating antibodies which are used to purify, detect
XX and quantify chitinases, e.g. for diagnosis of Gaucher's disease. The
XX nucleic acid sequence of the chitinase is also useful as a probe to
XX identify and isolate genomic DNA encoding chitinases or similar proteins,
XX or cells expressing them or to generate transgenic ('knockout') rodents.
XX It can also be used in hybridisation assays and to detect genetic
XX alterations in the chitinase gene related to disease. Agents that inhibit
XX this protein may be useful in treatment of Gaucher's disease and
XX rheumatoid arthritis, where overexpression of the protein can damage
XX the extracellular matrix. Chitinase also improves the activity of other
XX antifungal agents and may allow a reduction in the dose of such agents,
XX and thus of their side effects.
SQ Sequence 1656 BP; 365 A; 497 C; 447 G; 347 T; 0 other;

Alignment Scores:
Pred. No.: 9.53e-107 Length: 1656
Score: 1295.50 Matches: 234
Percent Similarity: 66.89% Conservative: 71
Best Local Similarity: 51.32% Mismatches: 130
Query Match: 51.97% Indels: 21
DB: 19 Gaps: 3

US-10-004-219B-9 (1-452) x AAV10436 (1-1656)
```


CC detection and quantification of chitinases. Polynucleotide fragments of
 CC the invention are useful as a source of probes and primers, and to
 CC express the proteins recombinantly. The chitinase fragments, when
 CC conjugated to antifungal compounds, are used to treat animals,
 CC especially humans, infected with chitin-containing parasites such as
 CC fungi. Fungal infection treated include candidiasis, aspergillosis,
 CC coccidioidomycosis, blastomycosis, paracoccidioidomycosis,
 CC mucormycosis, histoplasmosis, cryptococcosis, chromoblastomycosis,
 CC sporotrichosis, and dermatophytoses.
 CC Chitin can be degraded by the enzyme chitinase. Use of whole chitinase
 CC protein for treating infections, especially fungal infections, is
 CC problematic. In view of the increasing incidents of life-threatening
 CC fungal infection in e.g. immunocompromised individuals, there exists a
 CC need for identifying new compounds for treating fungal infection. The
 CC chitin-binding fragments of the present invention provide this need.

XX
 SQ Sequence 1656 BP; 365 A; 497 C; 447 G; 347 T; 0 other;

Alignment Scores:

Pred. No.: 9,53e-107 Length: 1656
 Score: 1295.50 Matches: 234
 Percent Similarity: 66.89% Conservative: 71
 Best Local Similarity: 51.32% Mismatches: 130
 Query Match: 51.97% Indels: 21
 DB: 20 Gaps: 3

US-10-004-219b-9 (1-452) x AAZ21848 (1-1656)

QY 3 LeuileCysThrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPheLysPro 22
 DB 96 CTGGTCTGCTACTTACCACCACTGGGCCCGGACGACGAGGGGGGAGGCTCGCTTCTCGGCC 155
 QY 23 AspAspIleAsnProCysLeuCysThrHisLeuileTyrAlaPheAlaGlyMetGlnAsn 42
 DB 156 AAGGACTTGGACCCCGAGCTTTGGACCCACCTACCTACCTGCTGGCATGACCAAC 215
 QY 43 AsnGluileThrIleGlnTrpAsnAspValThrLeuTyrLysAlaPheAsnAspLeu 62
 DB 216 CACCAGCTGAGCACCACTGAGTGGGAATGACGAGACTCTTACCAAGAGTTCAATGGCCTG 275
 QY 63 LysAsnArgAsnSerLysLeuLysThrLeuAlaIleGlyGlyTyrAsnPheGlyThr 82
 DB 276 AAGAAGATGAATCCCAAGCTGAAGACCCCTGTAGGCATCGGAGCTGGAAATTCAGCACT 335
 QY 83 AlaProPheThrMetValSerThrSerGlnAsnArgGlnThrPheIleThrSerVal 102
 DB 336 CAGAAGTTCACAGATATGTTAGCCAGCGGCCAACACACCTGTCAGACCTTTGTCACTGGCC 395
 QY 103 IleLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyrProGly 122
 DB 396 ATCAGGTTTCTGGCAATATACAGCTTTGACGGCTTGACCTTGACCTGGGAGTACCCAGGA 455
 QY 123 SerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValLysGluMetArg 142
 DB 456 AGCCAGGGGAGCCCTGCGCTAGACAGGAGCGCTTACACACCCCTGGTACAGACTTGGCC 515
 QY 143 GluAlaPheGluGluAlaIleGluSerAsnArgProArgLeuMetValThrAlaAla 162
 DB 516 AATGCTTCCACAGAGAACCCAGACCTCAGGAGGAGAACGCTTCTTCTGATGGCAGG 575
 QY 163 ValAlaGlyGlyIleSerAsnIleGlnAlaGlyTyrGluIleProGluLeuSerLysTyr 182
 DB 576 GTTCCAGCTGGGCGACACCTATGTGGATGTGTGATACGAGGTGGACAAATCGCCGAGAC 635
 QY 103 LeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThrGly 202
 DB 636 CTGGATTTTGTCAACCTTATGGCTACGACTTCCATGGCTCTTGGGAGAAGGTACGGGA 695
 QY 203 GluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeuAsnVal 222
 DB 696 CATACAGCCCCCTCTACAAAGAGGCGGAGAGAGAGTGGTGGCAGCGCCAGCTCACCGTG 755
 QY 223 AspTyrValMetAsnTyrTrpLysAsnGlyAlaProAlaGluLysLeuileValGly 242

DB 756 GATGCTGCTGTGCAACAGTGGCTGTCAGAAAGGGGAGCCCTGCCAGCAAGCTGATCCTGGC 815
 QY 243 PheProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIleGlyAla 262
 DB 816 ATGCTTACTACGAGCGCTCTTCTACACTGGCCCTCTCATCAGACACAGAGTGGGGGCC 875
 QY 263 ProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrpAlaTyr 282
 DB 876 CCAGCCACAGGCTGGGACCTCCAGGCCCTTCACCAAGGAGGAGGATGCTGGACCTAC 935
 QY 283 TyrGluileCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSerGlnGlu 302
 DB 936 TATGAAGTCTGCTCTGG-----AAGGGGGCCCAACACAGAGAATCAGGATCAGAAG 989
 QY 303 ValProTyrAlaTyrLysAlaAsnGluTrpLeuGlyTyrAspAsnIleLysSerPheSer 322
 DB 990 GTGCCCTACATCTTCCGGGACACACGAGGGGTGGCTTTGATGATGTGGGAGCTTCAAA 1049
 QY 323 ValLysAlaGlnTrpLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTrpAlaIleAsp 342
 DB 1050 ACCAAGGTCACTATCTGAAGCAGAGAGGACTGGGGGGGGCCATGCTGGGCACTGGAC 1109
 QY 343 LeuAspAspPheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSerThrLeu 362
 DB 1110 TTAGATGACTTTGGCGGCTTCTCTGCAACACGAGGCGCATACCCCTCATCCAGACGCTA 1169
 QY 363 AsnLysAlaLeuGlyIleSerThrGluGlyCysThrAlaProAspVal----- 378
 DB 1170 CGGCAGGAACAGTGTCTTCCATACCTTCCAGCACCCAGAGCTTGAAGTTCACAAA 1229
 QY 379 -----ProSerGluProValThrThrProGlySerGlySerGlyGlyGlySer 395
 DB 1230 CCAGTCAAGCCCTCGAACCT-----GAG 1253
 QY 396 SerGlyGlySerSerGlySerGlyPheCysAlaAspLysAlaAspGlyLeuTyrPro 415
 DB 1254 CATGCCCCAGCCCTGGACAGACACAGCTTCTGCCAGGGCAAGCTGATGGGCTCTATCC 1313
 QY 416 ValAlaAspAspArgAsnAlaPheTrpGlnCysIleAsnGlyIleThrTyrGlnGlnHis 435
 DB 1314 AATCTCGGGAACGGTCCAGCTTCTACAGCTGTGAGCGGGGGGGCTGTTCAGCAAAAGC 1373
 QY 436 CysGlnAlaGlyLeuValPheAspThrSerCysAsnCysCysAsnTyr 451
 DB 1374 TGCCCGACAGGCTGTGTTCAGCAACTCCTGCAATGCTGCACCTGG 1421
 RESULT 12
 AAD03760
 ID AAD03760 standard; cDNA; 1656 BP.
 XX
 AC AAD03760;
 XX 19-JUN-2001 (first entry)
 DT
 DE Human chitinase cDNA from clone pMO-13B.
 DE
 XX Human; antifungal; chitinase; immunoglobulin; Ig; therapy;
 KW fungal infection; candidiasis; aspergillosis; coccidioidomycosis;
 KW blastomycosis; paracoccidioidomycosis; histoplasmosis; cryptococcosis;
 KW chromoblastomycosis; sporotrichosis; mucormycosis; dermatophytoses;
 KW clone pMO-13B; ss.
 XX
 OS Homo sapiens.
 XX
 Key Location/Qualifiers
 PH 27..1427
 CDS /tag= a
 FT /product= "Human chitinase from clone pMO-13B"
 FT /tag= b
 FT 27..89
 FT 90..1424
 FT /tag= c
 FT mat_peptide
 FT sig_peptide

FT XX /product= "Human mature chitinase from clone pMO-13B"

PN W0200123430-A2.

XX 05-APR-2001.

XX 28-SEP-2000; 2000W0-US26960.

XX 30-SEP-1999; 99US-0409918.

XX (ICOS-) ICOS CORP.

XX Allison DS, Dietsch GN, Gray PW, Shaw KD, Steiner BH;

PI WPI: 2001-266141/27.

DR P-PSDB; ABE00433.

DR Novel chitinase immunoglobulin fusion product, useful for treating

PT fungal infections and reducing the amount of a non-chitinase antifungal

PT agent needed for the treatment -

XX Claim 2; Page 34-36; 39pp; English.

PS The present invention relates to a chitinase immunoglobulin (Ig) fusion

CC product, comprising a human chitinase fused to at least a portion of an

CC immunoglobulin chain. The fusion product is useful for treating fungal

CC infections (mycoses) such as candidiasis, aspergillosis, blastomycosis,

CC coccidioidomycosis, paracoccidioidomycosis, histoplasmosis, mucormycosis,

CC cryptococcosis, chromoblastomycosis, sporotrichosis and dermatophytoses.

CC The fusion protein is useful for reducing the amount of non-chitinase

CC antifungal agent needed to exert an antifungal activity. The fusion

CC protein is also useful for preparing a medicament for the prophylactic

CC or therapeutic treatment of fungal infections. Chitinase immunoglobulin

CC fusion product has unexpectedly improved serum half-life and formulation

CC properties. The present sequence is human chitinase cDNA from clone

CC pMO-13B. Chitinase enzyme degrades chitin which is a homopolymer of

CC beta-(1,4)-linked N-acetylglucosamine residues.

XX Sequence 1656 BP; 365 A; 497 C; 447 G; 347 T; 0 other;

SQ

Alignment Scores:

Pred. No.:	9-53e-107	Length:	1656
Score:	1295.50	Matches:	234
Percent Similarity:	66.89%	Conservative:	71
Best Local Similarity:	51.32%	Mismatches:	130
Query Match:	51.97%	Indels:	21
DB:	22	Gaps:	3

US-10-004-219B-9 (1-452) x AAD03760 (1-1656)

QY	3	LeuileCysThrPheThrAsnTPalaGlnTyrAtgProGlyLeuGlySerPheLysPro	22
DB	96	CTGGTCTGTCTTACCACTGGGCGCCAGTACAGACAGGGGAGGCTCGCTTCCTGCCC	155
QY	23	AspAspLeuProCysLeuCysThrHisLeuileTyrAlaPheAlaGlyMetGlnAsn	42
DB	156	AAGGACTTGGACCCAGCCTTTGACCCACCTCATCTACGCTTGTGCGATGACCAAC	215
QY	43	AsnGluileThrThrIleGlnTrpAsnAspValThrLeuTyrLysAlaPheAsnAspLeu	62
DB	216	CACCAGCTGAGCACCCTCAGTGGATGATGAGAGCTCTTACGAGGTTCAATGGCCTG	275
QY	63	LysAsnArgAsnSerLysLeuLysThrLeuLeuAlaileGlyGlyTyrAsnPheGlyThr	82
DB	276	AGAAGATGATCCCAAGCTGAGACCTTGTAGCCATCGGAGGCTGGAAATTCAGCACT	335
QY	83	AlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheIleThrSerVal	102
DB	336	CAGAAAGTTCACAGATATGTTAGTCCAGCGGCCAACACCGTCAGACCTTTGTCAACTCGGC	395
QY	103	IleLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGlnTyrProGly	122
DB	396	ATCAGGTTTTCGGCAATACAGCTTTGACGGCTTGACCTTGACTGGGAGTACCAGGA	455

QY	123	SerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValLysGluMetArg	142
DB	456	AGCCAGGGAGCCCTGCGTAGACAGGAGCGCTTCACACCTCGTACAGACTTGGCC	515
QY	143	GluAlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThrAlaAla	162
DB	516	AATGCTTCCAGCAGGAAGCCAGACCTCAGGAGGAACGCTTCTTCTGAGTCAGCG	575
QY	163	ValAlaGlyGlyIleSerAsnIleGlnAlaGlyTyrGluIleProGluLeuSerLysTyr	182
DB	576	GTTCAGCTGGGACAGCCTATGTGGATGCTGATACGAGGTGGCAAAATGCCAGAAC	635
QY	183	LeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTyrGluGlyTyrThrGly	202
DB	636	CTGGATTTGTCAACCTTATGGCTTACGACTTCCATGCTCTTGGAGAGTCAAGGGA	695
QY	203	GluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeuAsnVal	222
DB	696	CATAACAGCCCTCTACAAGAGGCAAGAGAGTGTGTGAGCAGCAGCCAGCTCAACGTG	755
QY	223	AspTyrValMetAsnTyrTrpLysAsnAsnGlyAlaProAlaGluLysLeuIleValGly	242
DB	756	GATGCTGCTGTGCAACAGTGGCTGCAGAGGGGACCCCTGCCAGCAAGCTGATCTTGGC	815
QY	243	PheProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIleGlyAla	262
DB	816	ATGCTTACTACGAGCGCTCTTCACTGGCTCTCATCAGACACCCAGAGTGGGGCC	875
QY	263	ProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrpAlaTyr	282
DB	876	CCAGCCACAGGGCTCGGACCTCCAGGCCCTTCAACAGGAAGGAGGATGTGCTGCTAC	935
QY	283	TyrGluileCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSerGlnGlu	302
DB	936	TATGAAGTCTGCTCTGCG-----AAGGGGGCCACCAACAGAGAAATCCAGATCAGAG	989
QY	303	ValProTyrAlaTyrLysAlaAsnGluTrpLeuGlyTyrAspAsnIleLysSerPheSer	322
DB	990	GTGCTTACATCTTCCGGGACACACAGTGGTGGGCTTGTGATGTGGAGAGCTTCAAA	1049
QY	323	ValLysAlaGlnTrpLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTrpAlaIleAsp	342
DB	1050	ACCAGGTCAGCTATCTGAAGCAGAGGAGGACTGGGGGGGGCCATGTCTGGGCACTGGAC	1109
QY	343	LeuAspAspPheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSerThrLeu	362
DB	1110	TTAGATGACTTGGCGGCTTCTCTGCAACAGGGGCCGATACCCCTCATCCAGAGCGCTA	1169
QY	363	AsnLysAlaLeuGlyIleSerThrGluGlyCysThrAlaProAspVal-----	378
DB	1170	CGGACGAACTGAGCTTCCATCTTCCAGTCTTCCAGCACCCAGAGCTTGAAGTTCCAAA	1229
QY	379	-----ProSerGluProValThrThrProProGlySerGlySerGlyGlySer	395
DB	1230	CCAGGTCAGCCCTCTGAACCT-----	1253
QY	396	SerGlyGlySerSerGlyPheCysAlaAspLysAlaAspGlyLeuTyrPro	415
DB	1254	CATGCCCCCGCCCTGGACACAGCTTCTGCCAGGGCAAGCTGATGGGCTCTATCCC	1313
QY	416	ValAlaAspAspArgAsnAlaPheTrpGlnCysIleAsnGlyIleThrTyrGlnGlnHis	435
DB	1314	AATCTCGGGAACGCTCCAGCTTCTACAGCTGTGCGGGGGCGGCTGTCTCCAGCAAGC	1373
QY	436	CysGlnAlaGlyLeuValPheAspThrSerCysAsnCysCysAsnTrp	451
DB	1374	TGCCCCGACAGGCTGTGTTCAGCAACTCTCTGCAAAATGCTGCACCTGG	1421

RESULT 13

ABL57381

ID ABL57381 standard; cDNA; 1656 BP.

XX

ABL57381;
 12-AUG-2002 (first entry)
 Human chitinase cDNA clone MO-13B.
 Chitinase; enzyme; human; fungicide; antifungal; infection;
 candidiasis; aspergillosis; coccidioidomycosis; blastomycosis;
 paracoccidioidomycosis; histoplasmosis; cryptococcosis;
 chromoblastomycosis; sporotrichosis; mucormycosis; dermatophytosis;
 Pneumocystis; gene; ss.
 Homo sapiens.
 Key Location/Qualifiers
 CDS 27..1427 /tag= a
 /product= "Chitinase"
 sig_peptide 27..89 /tag= b
 mat_peptide 90..1424 /tag= c
 US6372212-B1.
 16-APR-2002. 97US-0877599.
 16-JUN-1997; 96US-0663618.
 14-JUN-1996; 96US-0663618.
 (ICOS-) ICOS CORP.
 Gray PW;
 WPI; 2002-442449/47.
 P-PSDB; ABB76292.
 Co-administering chitinase to improve the effectiveness of fungicidal
 drugs e.g. amphotericin B or itraconazole, useful for treating fungal
 infections e.g. candidiasis, coccidioidomycosis and blastomycosis -
 Example 1; Column 31-34; 26pp; English.
 The present sequence is that of cDNA clone MO-13B (ATCC 98078)
 encoding human chitinase (see ABB76292). The clone was isolated
 from a cDNA library prepared from peripheral blood monocyte-derived
 macrophages using a probe based on human chitinase clone MO-218
 (see ABL57380). The 2 clones contain a single nucleotide
 difference in the coding region, causing amino acid position 81 of
 the mature protein to be either serine or glycine. Northern blots
 showed highest chitinase gene expression in lung and ovary tissues.
 Expression in lung is consistent with a protective role against
 pathogenic organisms that contain chitin. The invention provides
 human chitinase polynucleotides and polypeptides, and materials and
 methods for the recombinant production of human chitinase products,
 which are expected to be useful as products for treating fungal
 infections or for the development of such products. Human
 chitinase has a synergistic effect on the actions of other
 fungicides. It can be administered to improve the antifungal
 activity of a non-chitinase antifungal agent, especially
 amphotericin B or itraconazole, in the treatment of a fungal
 infection such as candidiasis, aspergillosis, coccidioidomycosis,
 blastomycosis, paracoccidioidomycosis, histoplasmosis,
 cryptococcosis, chromoblastomycosis, sporotrichosis, mucormycosis,
 dermatophycosis and Pneumocystis infections (all claimed). In
 particular, the fungal infection involves Candida, Aspergillus
 and/or Cryptococcus spp., whose growth is not effectively
 inhibited by contact with human chitinase alone.
 Sequence 1656 BP; 365 A; 497 C; 447 G; 347 T; 0 other;
 Alignment Scores:

Pred. No.:	9.53e-107	Length:	1656
Score:	1295.50	Matches:	234
Percent Similarity:	66.89%	Conservative:	71
Best Local Similarity:	51.32%	Mismatches:	130
Query Match:	51.97%	Indels:	21
DB:	24	Gaps:	3

US-10-004-219B-9 (1-452) x ABL57381 (1-1656)

QY	3	LeuileCysTyrPheThrAsnTyrAlaGlnTyrArgProGlyLeuGlySerPheLysPro	22
DB	96	CTGGTCTGTACTTCCACCACTGGGCCAGTACAGACAGGGGGAGGCTCGCTTCTGTCGCC	155
QY	23	AspAspIleAsnProCysLeuCysThrHisLeuileTyrAlaPheAlaGlyMetGlnAsn	42
DB	156	AAGGACTTGGACCCAGCCCTTTGACCCACCTCATCTACGGCTTCCTGGGATGACCAAC	215
QY	43	AsnGluileThrThrIleGluTyrAsnAspValThrLeuTyrLysAlaPheAsnAspLeu	62
DB	216	CACCAGCTGAGCACCCACTGAGTGAATGACGAGACTCTTACAGGAGTTCATATGGCCTG	275
QY	63	LysAsnArgAsnSerLysLeuLysThrLeuLeuAlaIleGlyGlyTyrAsnPheGlyThr	82
DB	276	AAGAAGATGAATCCCAAGCTGAAGACCTGTAGCCATCGGAGGCTGGAATTTACGCACT	335
QY	83	AlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheIleThrSerVal	102
DB	336	CAGAAGTTCACAGATATGTAGCCAGCCGCAACACCTGTGACACCTTTGTCACTCGGCC	395
QY	103	IleLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTyrGlyTyrProGly	122
DB	396	ATCAGGTTTCTGGCAATACAGCTTTTCAGCGCTTGACCTTGACCTGGGAGTACCCAGGA	455
QY	123	SerArgGlySerProGlnAspLysHisLeupheThrValLeuValLysGluMetArg	142
DB	456	AGCCAGGGGAGCCCTGCGGTAGACAAAGGAGGCGCTTCAACACCTCGTACGACTTGGCC	515
QY	143	GluAlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThrAlaAla	162
DB	516	AATGCTTTCAGCAGGAGAGCCAGACCTCAGGAGGAGAACGCTTCTTCTGAGTGCAGGG	575
QY	163	ValAlaGlyGlyIleSerAsnIleGlnAlaGlyTyrGluileProGluLeuSerLysTyr	182
DB	576	GTTCAGCTGGGAGACCTATGTGATCTGTGATCTGATACGAGGTGGACAAATCCCCAGAAC	635
QY	183	LeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTyrGluGlyTyrThrGly	202
DB	636	CTGGATTTTGTCAACCTTATGGCTTACGACTTCCATGGCTCTTGGGAGAGGTTCAGGGA	695
QY	203	GluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeuAsnVal	222
DB	696	CATAACACCCCTCTACAGAGGCAAGAGAGTGTGTGACAGCAGCAGCCTCAACGTG	755
QY	223	AspTyrValMetAsnTyrTrpLysAsnAsnGlyAlaProAlaGluLysLeuileValGly	242
DB	756	GATGCTGTGTGCAACACTGGCTGCAGAGGGGAGCCCTCGCCAGCAGCTGATCTTGGC	815
QY	243	PheProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIleGlyAla	262
DB	816	ATGCTACTACGAGCGCTCTTACACTGGCTCTCATCAGACACCCAGAGTGGGGGCC	875
QY	263	ProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrpAlaTyr	282
DB	876	CCAGCCACAGGCTTGGGCACTCCAGGCCCTTCCACAGGAGGAGGAGGATGTCGCTTAC	935
QY	283	TyrGluileCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSerGlnGlu	302
DB	936	TATGAAGTCTGCTCTGG-----AAGGGGGCCACCAACAGAGAATCCAGGATCAGAAG	989
QY	303	ValProTyrAlaTyrLysAlaAsnGluTrpLeuGlyTyrAspAsnIleLysSerPheSer	322
DB	990	GTGCCCTACATCTTCGGGCAACACGATGGGTGGGCTTTGTGATGTGGAGAGCTTCAA	1049


```
db 913 ATGCCTACCTACGGACGCTCTCTACACTGGCTCTCTATCAGACACACAGAGTGGGGCC 972
Qy 263 ProThrSerGlyAspGlyProAlaGlyAlaThrArgGlnAlaGlyPheThrAlaTyr 282
Db 973 CAGCCACAGGCTGGCACTCCAGCCCTTCACCAAGGAAGGAGGATGTCGGCTTAC 1032
Qy 283 TyrGluileCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSerGlnGlu 302
Db 1033 TATGAAGTCTGCTCTGG-----AAGGGGGCCACCAACAGAGAATCCAGGATCAGAAG 1086
Qy 303 ValProTyrAlaTyrLysAlaGlnGluTrpLeuGlyTyrAspAsnIleLysSerPheSer 322
Db 1087 GTGCCCTACATCTCCGGGACACACAGTGGGTGGCTTTGATGATGGAGAGCTTCAAA 1146
Qy 323 ValLysAlaGlnTrpLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTrpAlaIleAsp 342
Db 1147 ACCAGGTCAGTATCTGAGCAGAGGACTGGGGGGCCATGCTCTGGCACTGGAC 1206
Qy 343 LeuAspAspPheThrGlySerPheCysAspGlnGlyLysPhePheProLeuThrSerThrLeu 362
Db 1207 TTAGATGACTTTTGGCGCTTCTCTGCAACACAGGCGCGATACCCCTCATCCAGAGCTA 1266
Qy 363 AsnLysAlaLeuGlyLysSerThrGluGlyCysThrAlaProAspVal----- 378
Db 1267 CGGCAGGAACCTAGCTTCTCATACTTGCCTTCAGGCACCCAGAGTTGAAGTTCAAAA 1326
Qy 379 -----ProSerGluProValThrThrProProGlySerGlySerGlyGlySer 395
Db 1327 CAGGTGAGCCCTCTGAACCT-----GAG 1350
Qy 396 SerGlyGlySerGlyGlySerGlyPheCysAlaAspLysAlaAspGlyLeuTyrPro 415
Db 1351 CATGGCCCGAGCCCTGGACAGACACAGTCTCTGCCAGGGCAAGCTGATGGGCTCTATCCC 1410
Qy 416 ValAlaAspAspArgAsnAlaPheTrpGlnCysIleAsnGlyIleThrTyrGlnGlnHis 435
Db 1411 AATCTCGGGACGGTCCAGCTTCTACAGCTGTGCAGCGGGCGGCTGTTCAGCAAAAGC 1470
Qy 436 CysGlnAlaGlyLeuValPheAspThrSerCysAsnCysCysAsnTrp 451
Db 1471 TGCCGACAGGCGCTGTGTTCAGCAACTCTGCAANTGTCACCTGG 1518

RESULT 15
ID AAT50834
XX AAT50834 standard; cDNA; 1713 BP.
AC AAT50834;
XX
XX 24-MAR-1997 (first entry)
XX Human chitinase cDNA clone chi.39.
XX Chitinase; chitotriosidase; chitin; infectious disease;
KW gene therapy; vaccine; lysosomal lipidoses; Gaucher's disease;
KW leishmaniasis; sarcoidosis; X-linked adrenoleukodystrophy;
KW multiple sclerosis; drug delivery; cosmetics; food; ss.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 13..1176
XX FT /*tag= a
XX FT sig_peptide 13..75
XX FT /*tag= b
XX FT mat_peptide 76..1173
XX FT /*tag= c
XX
XX WO9640940-A2.
XX
XX 19-DEC-1996.
XX
XX 06-JUN-1996; 96WO-NL00225.
```

```
XX 07-JUN-1995; 95US-0486839.
PR (UNAM ) UNIV AMSTERDAM.
XX
XX Aerts JMG;
XX
XX WPI; 1997-118698/11.
XX P-PSDB; AAW08585.
XX
XX New human chitinase - used to treat or prevent infection by
XX chitin-contg. pathogens, in diagnosis and as additives to cosmetics,
XX foods, implants etc.
XX
XX Claim 2; Page 42-43; 58pp; English.
XX
XX A cDNA clone (AAT50834), designated chi.39, codes for a 39 kDa human
XX chitinase (AAW08584) that is stable to many proteases, active at pH 3-
XX 8 and up to 50 deg, and stable in the circulation. Clones chi.39
XX and chi.50 (see also AAT50833) were isolated from a human macrophage
XX cDNA library using as probe a partial clone obtd. using primers
XX (see also AAT50835-36) based on a chitotriosidase purified from a
XX type 1 Gaucher disease patient. The 2 cDNA clones are the result
XX of alternative splicing of RNA. Chitinase nucleic acid can be used
XX for large-scale prodn. of recombinant human chitinases, or can be
XX incorporated into a gene therapy vector to treat or prevent
XX infection by chitin-contg. pathogens.
XX
XX Sequence 1713 BP; 380 A; 503 C; 465 G; 365 T; 0 other;
SQ

Alignment Scores:
Pred. No.: 7,12e-106 Length: 1713
Score: 1286.00 Matches: 240
Percent Similarity: 66.52% Conservative: 72
Best Local Similarity: 51.17% Mismatches: 134
Query Match: 18 Indels: 24
DB: 5

US-10-004-219B-9 (1-452) x AAT50834 (1-1713)
Qy 3 LeuIleCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPheLysPro 22
Db 82 CTGGTCTGCTACTTCCACCAACTGGGCGCCAGTACAGACAGGGGGAGGCTCGTCTCTGCC 141
Qy 23 AspAspIleAsnProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyMetGlnAsn 42
Db 142 AAGGACTTGGACCCCGCCGCTTTGACACCCACTCATCTACGCTTCGCTGGCATGCCAAC 201
Qy 43 AsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrLysAlaPheAsnAspLeu 62
Db 202 CACCAGCTGAGCACCACCTGAGTGGAAATGACGAGACTCTTACCAGGAGTTCAATGGCCTG 261
Qy 63 LysAsnArgAsnSerLysLeuLysThrLeuLeuAlaIleGlyTyrAsnPheGlyThr 82
Db 262 AAGAAGATGAATCCCAAGCTGAAGACCTGTGTAGCATCGGAGGCTGGAATTTCCGCACT 321
Qy 83 AlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheIleThrSerVal 102
Db 322 CAGAAGTTCCACAGATATGTTAGCCACGCGCCAAACCCCTTGTCAACTCGGCC 381
Qy 103 IleLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyrProGly 122
Db 382 ATCAGTTTCTGGCAATAACAGCTTTGACCGGCTTGACCTTGAGTGGAGTACCAGGA 441
Qy 123 SerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValLysGluMetArg 142
Db 442 AGCCAGGGGAGCCCTGCGGTAGACAAAGGAGCGCTTCAACACCTCGTACAGACTTGGCC 501
Qy 143 GluAlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThrAlaAla 162
Db 502 AATGCTTCCAGCAGGAAGCCACCTCAGCGGAAGAACGCCCTTCTTCTGAGTGACGCG 561
Qy 163 ValAlaGlyGlyIleSerAsnIleGlnAlaGlyTyrGluIleProGluLeuSerLysTyr 182
```

[illegible]

Search completed: June 29, 2003, 21:26:40
Job time : 280.065 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 29, 2003, 21:12:38 ; Search time 64.5237 Seconds
(without alignments)
2148.326 Million cell updates/sec

Title: US-10-004-219B-9

Perfect score: 2493

Sequence: 1 YNLCYFTNMQAYRPLGSGF.....QQHQAGLVFTDSCNCCNWP 452

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+P2n.model -DEV=xlpl
-Q=/cgn2_1/USPTO_spool/US10004219/runat_24062003_160229_8184/app_query.fasta_1.1877
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rn1 -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptco -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10004219@cgn_1_1_108@runat_24062003_160229_8184 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:

1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCITUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result. No.	Score	Query Match	Length	ID	Description
1	1301.5	52.2	1636	US-09-039-198A-1	Sequence 1, Appli
2	1301.5	52.2	1636	US-08-877-599-1	Sequence 1, Appli
3	1301.5	52.2	1636	US-09-267-574-1	Sequence 1, Appli
4	1301.5	52.2	1643	US-08-486-839-3	Sequence 3, Appli
5	1301.5	52.2	1643	US-09-151-011-3	Sequence 3, Appli
6	1301.5	52.2	1643	US-09-343-623-3	Sequence 3, Appli
7	1295.5	52.0	1656	US-09-039-198A-3	Sequence 3, Appli
8	1295.5	52.0	1656	US-08-877-599-3	Sequence 3, Appli
9	1295.5	52.0	1656	US-09-267-574-3	Sequence 3, Appli
10	1286	51.6	1713	US-08-486-839-5	Sequence 5, Appli
11	1286	51.6	1713	US-09-151-011-5	Sequence 5, Appli
12	1286	51.6	1713	US-09-343-623-5	Sequence 5, Appli

13	971.5	39.0	1681	5	PCT-US94-07754-4	Sequence 4, Appli
14	966.5	38.8	1681	2	US-08-581-527-4	Sequence 4, Appli
15	950.5	38.1	1433	1	US-08-694-915-1	Sequence 1, Appli
16	950.5	38.1	1526	1	US-08-694-915-3	Sequence 3, Appli
17	863.5	34.6	2452	2	US-08-524-051-1	Sequence 1, Appli
18	802	32.2	1677	4	US-09-545-814-13	Sequence 13, Appli
19	802	32.2	1677	4	US-09-545-814-15	Sequence 15, Appli
20	802	32.2	1749	4	US-09-545-814-4	Sequence 4, Appli
21	802	32.2	1749	4	US-09-545-814-6	Sequence 6, Appli
22	802	32.2	2610	4	US-09-545-814-1	Sequence 3, Appli
23	802	32.2	2610	4	US-09-545-814-3	Sequence 3, Appli
24	799	32.0	1919	4	US-09-545-814-31	Sequence 31, Appli
25	729.5	29.3	1608	4	US-09-292-225-20	Sequence 20, Appli
26	729.5	29.3	1608	4	US-09-292-225-22	Sequence 22, Appli
27	729.5	29.3	1665	4	US-09-292-225-17	Sequence 17, Appli
28	729.5	29.3	1665	4	US-09-292-225-19	Sequence 19, Appli
29	729.5	29.3	1752	4	US-09-292-225-14	Sequence 14, Appli
30	729.5	29.3	1752	4	US-09-292-225-16	Sequence 16, Appli
31	729.5	29.3	1470	4	US-09-292-225-40	Sequence 40, Appli
32	722	29.0	1470	4	US-09-292-225-42	Sequence 42, Appli
33	722	29.0	1470	4	US-09-292-225-37	Sequence 37, Appli
34	722	29.0	1527	4	US-09-292-225-39	Sequence 39, Appli
35	722	29.0	1621	4	US-09-292-225-36	Sequence 36, Appli
36	722	29.0	1621	4	US-09-292-225-34	Sequence 34, Appli
37	722	29.0	1478	4	US-09-545-814-28	Sequence 28, Appli
38	667.5	26.8	1478	4	US-09-545-814-30	Sequence 30, Appli
39	667.5	26.8	1478	4	US-09-545-814-30	Sequence 30, Appli
40	470.5	18.9	1167	1	US-07-939-501A-6	Sequence 6, Appli
41	470.5	18.9	1167	4	US-08-448-398-10	Sequence 10, Appli
42	470.5	18.9	1320	1	US-07-939-501A-14	Sequence 14, Appli
43	470.5	18.9	1364	1	US-07-939-501A-13	Sequence 13, Appli
44	470.5	18.9	1405	1	US-07-939-501A-9	Sequence 9, Appli
45	461.5	18.5	1554	1	US-08-045-269C-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-09-039-198A-1

; Sequence 1, Application US/09039198A

; Patent No. 6200951

; GENERAL INFORMATION:

; APPLICANT: Gray, Patrick W.

; APPLICANT: Tjoelker, Larry W.

; TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS

; NUMBER OF SEQUENCES: 34

; CORRESPONDENCE ADDRESS:

; ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun

; STREET: 233 South Wacker Drive/6300 Sears Tower

; CITY: Chicago

; STATE: Illinois

; COUNTRY: United States of America

; ZIP: 60606-6402

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/039.198A

; FILING DATE:

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Rin-Laures, Li-Hsien

; REGISTRATION NUMBER: 33,547

; REFERENCE/DOCKET NUMBER: 27866/34391

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (312) 474-6300

; TELEFAX: (312) 474-0448

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1636 base pairs

; TYPE: nucleic acid

STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 2..1399
 FEATURE:
 NAME/KEY: mat peptide
 LOCATION: 65..1399
 US-09-039-198A-1

Alignment Scores:

Pred. No.: 6,7e-121 Length: 1636
 Score: 1301.50 Matches: 235
 Percent Similarity: 67.11% Conservatives: 71
 Best Local Similarity: 51.54% Mismatches: 129
 Query Match: 52.21% Indels: 21
 DB: 4 Gaps: 3

US-10-004-219B-9 (1-452) x US-09-039-198A-1 (1-1636)

```

QY 3 LeuIleCysThrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPheLysPro 22
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
71 CTGGTCTGTACTTCCACCACTGGGCCAGTACAGACAGGGGAGGCTCGCTTCTCGGCC 130

QY 23 AspAspIleAsnProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyMetGlnAsn 42
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
131 AAGGACTTGGACCCCGCCGCTTTGCACCCACCTCACTACGCTTCTCGCCATGACCAAC 190

QY 43 AsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrLysAlaPheAsnAspLeu 62
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
191 CACCACTGACGACCACTGAGTGAATGACGAGACTCTTACAGGAGTTCAATGGCCTG 250

QY 63 LysAsnArgAsnSerLysLeuLysThrLeuLeuAlaIleGlyThrAsnPheGlyThr 82
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
251 AAGAAGATCAATCCCAAGCTGAAGACCTGTAGCCATCGAGGCTGGAATTTCCGCACT 310

QY 83 AlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheIleThrSerVal 102
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
311 CAGAGTTTACAGATATGTAGCCAGGCGCAACACCGTCAGACCTTTGTCACTCGGCC 370

QY 103 IleLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyrProGly 122
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
371 ATCAGGTTTCTCGCAATACAGCTTTGACGCGCTTGACCTTGAGTGGAGTACCAGGA 430

QY 123 SerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValLysGluMetArg 142
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
431 AGCCAGGGGAGCCCTGCGCTAGACAGAGGAGCGCTTCAACACCTGTGTACAGACTTGGCC 490

QY 143 GluAlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThrAlaAla 162
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
491 ANTGCCTTCCAGAGGAGGAGCCAGACCTCAGGAGGAGGAGCGCTTCTGTAGTGCAGCG 550

QY 163 ValAlaGlyGlyIleSerAsnIleGlnAlaGlyTyrGluIleProGluLeuSerLysTyr 182
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
551 GTTCCAGCTGGGACAGACCTATGTGATGCTGTGATGAGTGGACAAATCCGCCAGAAC 610

QY 183 LeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThrGly 202
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
611 CTGGATTTTGTCAACCTTATGCTTACGACTTCCATGGCTTCTTGGAGAGTCAAGCGGA 670

QY 203 GluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeuAsnVal 222
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
671 CATACAGCCCTTCTACAGAGGAGGAGGAGAGAGTGTGAGCAGCCAGCTCAACGTG 730

QY 223 AspTyrValMetAsnTyrTrpLysAsnAsnGlyValProAlaGluLysLeuIleValGly 242
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
731 GATGCTGTGTCACACAGTGGCTGCGAGAGGAGGAGCCCTGCGCAGCAAGCTGATCTTGGC 790

QY 243 PheProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIleGlyAla 262
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
791 ATGCTTACCTACGAGGAGCTCTTCACTGAGCTGCTCTCATCAGACCAAGCTGGGGGCC 850

```

```

QY 263 ProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrpAlaTyr 282
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
851 CCAGCCACAGGCTCTGGCACTCCAGGCCCTTCCACCAAGGAGGAGGATGCTGGCCTAC 910

QY 283 TyrGluIleCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSerGlnGlu 302
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
911 TATGAAGTCTGCTCTGG-----AAGGGGGCCACCAACACAGAGAATCCAGGATCAGAAG 964

QY 303 ValProTyrAlaTyrLysAlaAsnGluTrpLeuGlyTyrAspAsnIleLysPheSer 322
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
965 GTGCCCTACATCTCCGGGCAACACGAGGGGGCTTTGATGATGTGGAGAGCTTCAAA 1024

QY 323 ValLysAlaGlnTrpLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTrpAlaIleAsp 342
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1025 ACCAAGTCACTATCTGACGAGAGGAGGACTGGGGGGGCGCATGTCTGGGCACTGGAC 1084

QY 343 LeuAspAspPheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSerThrLeu 362
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1085 TTAGATGACTTTGCGGCTTCTCTCTGCAACACGAGGCGCATACCCCTCATCCAGACGCTA 1144

QY 363 AsnLysAlaLeuGlyIleSerThrGluGlyCysThrAlaProAspVal----- 378
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1145 CGGCAGGAACACTGAGTCTTCCATACCTTCCAGGACCCAGAGCTTGAAGTCCAAAA 1204

QY 379 -----ProSerGluProValThrThrProGlySerGlySerGlyGlyGlySer 395
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1205 CCAGTCAAGCCCTCTGAACCT-----GAG 1228

QY 396 SerGlyGlySerSerGlyGlySerGlyPheCysAlaAspLysAlaAspGlyLeuTyrPro 415
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1229 CATGCCCCAGCCCTGGACAAGACACACAGTCTTCCAGGCGCAAGCTGATGGGCTCTATCCC 1288

QY 416 ValAlaAspAspArgAsnAlaPheTrpGlnCysIleAsnGlyIleThrTyrGlnGlnHis 435
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1289 AATCTCGGAGAGCGTCCAGCTTCTACAGCTGTGCGGGGGCGGCTGTTCAGCAAGC 1348

QY 436 CysGlnAlaGlyLeuValPheAspThrSerCysAsnCysCysAsnTrp 451
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1349 TGCCCGCAGGCGCTGTCTTTCAGCAACTCTCTGCAAAATGCTGCACCTGG 1396

```

RESULT 2

US-08-877-599-1
 ; Sequence 1, Application US/08877599
 ; Patent No. 6372212
 ; GENERAL INFORMATION:
 ; APPLICANT: Gray, Patrick W.
 ; TITLE OF INVENTION: Chitinase Materials and Methods
 ; NUMBER OF SEQUENCES: 17
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
 ; STREET: 6300 Sears Tower, 233 South Wacker Drive
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: United States of America
 ; ZIP: 60606-6402
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/877,599
 ; FILING DATE:
 ; CLASSIFICATION: 514
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/663,618
 ; FILING DATE: 14-JUN-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Rin-Laures Li-Hsien
 ; REGISTRATION NUMBER: 33,547
 ; REFERENCE/DOCKET NUMBER: 27866/33994
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 312/474-6300

TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1636 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 2..1399
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 65..1399
US-08-877-599-1

Alignment Scores:

Pred. No.: 6, 7e-121 Length: 1636
Score: 1301.50 Matches: 235
Percent Similarity: 67.11% Conservative: 71
Best Local Similarity: 51.54% Mismatches: 129
Query Match: 52.21% Indels: 21
DB: 4 Gaps: 3

US-10-004-219B-9 (1-452) x US-08-877-599-1 (1-1636)

QY 3 LeuileCysThrPheAsnTrpAlaGlnTrpArgProGlyLeuGlySerPheLysPro 22
DB 71 CTGGTCTGCTACTTCACCAACTGGGCCAGTACACAGAGGGAGGCTCGCTTCCTGGCC 130
QY 23 AspAspIleAsnProCysLeuCysThrHisLeuileTyAlaPheAlaGlyMetGlnAsn 42
DB 131 AAGGACTTGGACCCAGGCTTTGACCCACCTCATCTAGCCCTTCGCTGGCATGACCAAC 190
QY 43 AsnGluileThrThrIleGluTrpAsnAspValThrLeuTyLysAlaPheAsnAspLeu 62
DB 191 CACCAGCTGAGCACCACTGAGTGAATGACGAGACTCTTACAGAGTTCAATGGCCTG 250
QY 63 LysAsnArgAsnSerLysLeuLysThrLeuLeuAlaIleGlyLysTrpAsnPheGlyThr 82
DB 251 AAGAAGATGAATCCCAAGCTGAAGACCTGTAGCCATCGAGGCTGGAATTCGGCACT 310
QY 83 AlaProPheThrMetValSerThrSerGlnAsnArgGlnThrPheIleThrSerVal 102
DB 311 CAGNAGTTTACAGATATGTTAGCCAGGCCACCAACACCTTCAGACCTTCTCAACTGGCC 370
QY 103 IleLysPheLeuArgGlnTyThrGlyPheAspGlyLeuAspLeuAspTrpGluTyThrProGly 122
DB 371 ATCAGGTTTCTCGCAAAATACACAGCTTTGACGGCCTTGACCTTGAGTGGGAGTACCAGGA 430
QY 123 SerArgGlySerProGlnAspLysHisLeuPheThrValLeuValLysGluMetArg 142
DB 431 AGCCAGGGAGGCGCTCCCTGACAGAGGAGCGCTTCACAACTCTGGTACAGACTTGGCC 490
QY 143 GluAlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThrAlaAla 162
DB 491 ATGCTCTTCACAGGAGGCCAGACCTCAGGGAAGGAGCGCTTCTTGAGTGCAGG 550
QY 163 ValAlaGlyGlyLysSerAsnIleGlnAlaGlyTyThrGluileProGluLeuSerLysTy 182
DB 551 GTTCCAGCTGGGCAGACCTATGCGATGCTGATGACGAGTGGACAAATGCCAGAAC 610
QY 183 LeuAspPheIleHisValMetThrTyThrAspLeuHisGlySerTrpGluTyThrGly 202
DB 611 CTGGATTTTGTCAACCTTATGCCCTACGACTTCCATGGCTTTGGGAGAGTCAAGGGA 670
QY 203 GluAsnSerProLeuTyLysTyThrProThrGluThrGlySerAsnAlaTyThrLeuAsnVal 222
DB 671 CATACAGCCCTCTACAGAGGCCAGNAGAGAGTGGTGCAGCCAGCTTCAACGTG 730
QY 223 AspTyValMetAsnTyThrTyLysAsnGlyAlaProAlaGluLysLeuileValGly 242

DB 731 GATGCTGCTGTGCAACAGTGGCTGCGAGAAGGGGCCCTCCAGCAAGACTGATCCTTGGC 790
QY 243 PheProGluTyThrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIleGlyAla 262
DB 791 ATGCTTACCTACGACGCTCTTACACTGGCTCTCTCATCAGACACCAGAGTGGGGCC 850
QY 263 ProThrSerGlyAspGlyProAlaGlyAlaTyThrArgGlnAlaGlyPheTrpAlaTy 282
DB 851 CCAGCCACAGGCTGTCGACCTCCAGGCCCTTCCACAGGAAGGAGGATGTCGCTGCTAC 910
QY 283 TyrGluileCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSerGlnGlu 302
DB 911 TATGAAGTCTGCTCTCTGG-----AAGGGGGCCCAACACAGAAATCCAGATCAGAA 964
QY 303 ValProTyAlaTyLysAlaAsnGluTrpLeuGlyTyThrAspAsnIleLysSerPheSer 322
DB 965 GTGCTTACATCTTCCGGGACACAGTGGTGGGCTTTGATGATGTGGAGAGCTTCAA 1024
QY 323 ValLysAlaGlnTrpLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTrpAlaIleAsp 342
DB 1025 ACCAAGGTGACGTATCTGAAGCAGAGGACTGGGGGGGCGCATGGTCTGGGCACCTGAC 1084
QY 343 LeuAspAspPheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSerThrLeu 362
DB 1085 TTAGATGACTTTGGCGGCTTCTCTGCAACAGGCGCGATACCCCTCATCCAGACGCTA 1144
QY 363 AsnLysAlaLeuGlyIleSerThrGluGlyCysThrAlaProAspVal----- 378
DB 1145 CGGAGGAGTGTAGTCTTCCATCTTGGCTTCAGGACCCCGAGAGTTGAAGTTCACAAA 1204
QY 379 -----ProSerGluProValThrThrProGlySerGlySerGlyGlySer 395
DB 1205 CCAGTTCAGCCCTCTGAACCT-----GAG 1228
QY 396 SerGlyGlySerSerGlyGlySerGlyPheCysAlaAspLysAlaAspGlyLeuTyPro 415
DB 1229 CATGGCCCGACCCCTGGCAAGACACAGCTTCTGCCAGGCGCAAGCTGATGGGCTCTATCCC 1288
QY 416 ValAlaAspAspArgAsnAlaPheTrpGlnCysIleAsnGlyIleThrTyGlnGlnHis 435
DB 1289 AATCTCGGGAACGGTCCAGCTTTACAGCTGTGCGAGCGGGCGGCTGTTCCAGCAAGC 1348
QY 436 CysGlnAlaGlyLeuValPheAspThrSerCysAsnCysCysAsnTrp 451
DB 1349 TGCCGACAGGCGCTGTGTTCAGCAACTCTCTGCAAAATGCTGCACCTGG 1396

RESULT 3

US-09-267-574-1
; Sequence 1, Application US/09267574
; Patent No. 6399571
; GENERAL INFORMATION:
; APPLICANT: Gray, Patrick W.
; APPLICANT: Tjoelker, Larry W.
; TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
; FILE REFERENCE: 27866/35407
; CURRENT APPLICATION NUMBER: US/09/267,574
; CURRENT FILING DATE: 1999-03-12
; EARLIER APPLICATION NUMBER: 09/039,198
; EARLIER FILING DATE: 1998-03-12
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1636
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(1399)
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (65)..(1399)
US-09-267-574-1

Alignment Scores:

Pred. No.: 6.7e-121 Length: 1636
Score: 1301.50 Matches: 235
Percent Similarity: 67.11% Conservative: 71
Best Local Similarity: 51.54% Mismatches: 129
Query Match: 52.21% Indels: 21
DB: 4 Gaps: 3

US-10-004-219B-9 (1-452) x US-09-267-574-1 (1-1636)

```
QY 3 LeuileCysThrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPheLysPro 22
Db 71 CTGGTCTGTACTTACCAACTGGGCCAGTACAGACAGGGGAGGCTCGCTTCTGCGCC 130
QY 23 AspAspIleAsnProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyMetGlnAsn 42
Db 131 AAGGACTTGGACCCAGCCTTTGACCCACCTCATCTAGCCCTTCTGCGATGACCAAC 190
QY 43 AsnGluileThrThrileGluTrpAsnAspValThrLeuTyrLysAlaPheAsnAspLeu 62
Db 191 CACCAGCTGAGCACCAGTGTGATGATGAGAGACTCTCTACAGAGTTCAATGGCCTG 250
QY 63 LysAsnArgAsnSerLysLysLeuLeuAlaileGlyGlyTrpAsnPheGlyThr 82
Db 251 AAGAAGATGAATCCCAAGCTGAAGACCTGTGTAGCCATCGAGGCTGGAATTTGGCACT 310
QY 83 AlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheileThrSerVal 102
Db 311 CAGAAGTTTACAGATATGTAGTCCAGCGGCCACACCGTCAGACCTTTGTCAACTCGGCC 370
QY 103 IleLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyrProGly 122
Db 371 ATCAGGTTTCTGGCAAAATACAGCTTTGAGCGCTTGACCTTGACCTGGAGTACCAGGA 430
QY 123 SerArgLysSerProGlnAspLysHisLeuPheThrValLeuValLysGluMetArg 142
Db 431 AGCCAGGGAGCCCTCCCTAGACAAAGAGCGCTTACCAACCTGTGTACAGGACTGGCC 490
QY 143 GluAlaPheGluGlnGluAlaileGluSerAsnArgProArgLeuMetValThrAlaAla 162
Db 491 ATGCTCTTCCAGCAGGAAGCCAGACTCAGGAAGGAACCGCTTCTGTAGTGCAGCG 550
QY 163 ValAlaGlyGlyIleSerAsnIleGlnAlaGlyTyrGluIleProGluLeuSerLysTyr 182
Db 551 GTTCCAGCTGGGCAGACCTATGTGATGTGTGATACAGGTGACAAATGCCCAGAAC 610
QY 183 LeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThrGly 202
Db 611 CTGGATTTTGTCAACCTTATGGCTTACGACTTCCATGGCTTCTGGGAGAAGGTACGGGA 670
QY 203 GluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeuAsnVal 222
Db 671 CATACAGCCCCCTTACAAAGAGGCAAGAGAGTGGTGACAGCCAGCCTCAACGTG 730
QY 223 AspTyrValMetAsnTyrTrpLysAsnAnglyAlaProAlaGluLysLeuIleValGly 242
Db 731 GATGCTGTGTGCAACAGTGGCTGCAGAGGGGACCCCTGCCAGCAGCTGATCTTGGC 790
QY 243 PheProGluTyrGlyHisThrPheileLeuArgAsnProSerAspAsnGlyIleGlyAla 262
Db 791 ATGCTTACCTACGAGGCTTCTTACACTGGCTTCTCATCATCACACACAGAGTGGGGCC 850
QY 263 ProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrpAlaTyr 282
Db 851 CAGCCACAGGGTCTGGCACTCAGGCCCTTCCACCAAGAGAGGAGGATGTGGCCTAC 910
QY 283 TyrGluileCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSerGlnGlu 302
Db 911 TATGAAGTGTCTCTGG-----AAGGGGGCCACCAACAGAGATCCAGGATCAGAG 964
QY 303 ValProTyrAlaTyrLysAlaGlnGluTrpLeuGlyTyrAspAsnIleLysSerPheSer 322
Db 965 GTGCCCTATCTTCGGGACCAACCAAGTGGTGGGCTTTGATGATGTGGAGAGCTTCAA 1024
```

```
QY 323 ValLysAlaGlnTrpLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTrpAlaIleAsp 342
Db 1025 ACCAAGGTGAGTATCTGAAGCAGAGGAGCTGGCGCGGCCCATGGTCTGGGCATGGAC 1084
QY 343 LeuAspAspPheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSerThrLeu 362
Db 1085 TTAGATGACTTTGGCGGCTTCTCTGCAACAGGAGGCCCATACCCCTCATCCAGACGCTA 1144
QY 363 AsnLysAlaLeuGlyLysSerThrGluGlyCysThrAlaProAspVal----- 378
Db 1145 CGGCAGGAAGTCTGAGTCTTTCATCTTTCAGGACCCAGAGCTTGAAGTTTCAAAA 1204
QY 379 -----ProSerGluProValThrThrProGlySerGlySerGlyGlyGlySer 395
Db 1205 CCAGTCTAGCCCTCTGAACCT-----GAG 1228
QY 396 SerGlyGlySerSerGlyGlySerGlyPheCysAlaAspLysAlaAspGlyLeuTyrPro 415
Db 1229 CATGCCCCAGCCCTGGACAAGACACAGCTTCTGCCAGGCGCAAGCTGATGGCTCTATCC 1288
QY 416 ValAlaAspAspArgAsnAlaPheTrpGlnCysIleAsnGlyIleThrTyrGlnGlnHis 435
Db 1289 AATCCTCGGGAACGCTCCAGCTTCTACAGCTGTGAGCGGGCGGCTGTTCAGCAAAAGC 1348
QY 436 CysGlnAlaGlyLeuValPheAspThrSerCysAsnCysCysAsnTrp 451
Db 1349 TGCCCGCAGCGCTGTGTGTGACAACTCTTGCAAAATGCTGCACCTGG 1396
```

RESULT 4

US-08-486-839-3
Sequence 3, Application US/08486839
Patent No. 5928928
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: A human chitinase, its recombinant
production, its use for decomposing chitin, its use
in therapy or prophylaxis against infection diseases.
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hoffmann & Baron
CITY: Jericho
STREET: 350 Jericho Turnpike
STATE: New York
COUNTRY: United States of America
ZIP: 11758
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486,839
FILING DATE: 07 - June - 1995
ATTORNEY/AGENT INFORMATION:
NAME: Baron, Ronald J.
REGISTRATION NUMBER: 29,281
REFERENCE/DOCKET NUMBER: 294-26
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 822-3550
TELEFAX: (516) 822-3582
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1643 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
US-08-486-839-3

Alignment Scores:

Pred. No.: 6,74e-121 Length: 1643
Score: 1301.50 Matches: 235
Percent Similarity: 67.11% Conservative: 71
Best Local Similarity: 51.54% Mismatches: 129
Query Match: 52.21% Indels: 21
DB: 2 Gaps: 3

US-10-004-219B-9 (1-452) x US-08-486-839-3 (1-1643)

QY 3 LeuileCysThrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPheLysPro 22
DB 82 CTGGTCTGCTACTTACCAACTGGGCCAGTAGACAGACAGGGGAGGCTCGCTTCTGCCCC 141
QY 23 AspAspIleAsnProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyMetGlnAsn 42
DB 142 AAGGACTGGACCCAGCCCTTGGACCCACCTCATCTAGCCCTTCGCTGGCATGCCAAC 201
QY 43 AenGluileThrThrIleGluTrpAsnAspValThrLeuTyrLysAlaPheAsnAspLeu 62
DB 202 CACCACTGAGCACCACCTGAGTGAATGACGAGACTCTCTACGAGAGTTCAATGGCCTG 261
QY 63 LysAsnArgAsnSerLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPheGlyThr 82
DB 262 AAGAAGATGAATCCCAAGCTGAAGACCCCTGTAGCCATCGGAGGCTGGAAATTCGCGACT 321
QY 83 AlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheIleThrSerVal 102
DB 322 CAGAAGTTCACAGATATGTCAGCCAGCCGACACACCTTGTCACTGGGCC 381
QY 103 IleLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyrProGly 122
DB 382 ATCAGGTTTCTGCGCAATACAGCTTTCAGCGCTTACCTTGTGAGTGGAGTACCCAGGA 441
QY 123 SerArgGlySerProGlnAspLysHisLeuPheThrValLeuValLysGluMetArg 142
DB 442 AGCCAGGGAGGCGCTGCGGTAGCAAGGAGCGCTTCAACACCTGTGTACAGACTTGGCC 501
QY 143 GluAlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThrAlaAla 162
DB 502 ANTGCCTTCACAGGAGGAGCCAGACCTCAGGGAGGAGCGCTTCTGTAGTGGAGG 561
QY 163 ValAlaGlyGlyIleSerAsnIleGlnAlaGlyTyrGluIleProGluLeuSerLysTyr 182
DB 562 GTTCCAGCTGGGAGACCTATGTGGATGCTGTGATACGAGTGGACAAATCGCCAGAAC 621
QY 183 LeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThrGly 202
DB 622 CTGGATTTTGTCAACCTTATGGCCTACGACTTCCATGGCTCTTGGGAGAGGTACGGGA 681
QY 203 GluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeuAsnVal 222
DB 682 CATACAGCCCTCTACAGAGGCGGAGAGAGAGTGGTGGAGCCAGCCCTCAACGTG 741
QY 223 AspTyrValMetAsnTyrTrpLysAsnAenGlyAlaProAlaGluLysLeuIleValGly 242
DB 742 GATGCTGCTGTGCAACAGTGGCTGCAAGAGGGGAGCCCTGCCAGCAAGCTGATCTTGGC 801
QY 243 PheProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIleGlyAla 262
DB 802 ATGCCTTACCTAGCGAGCGCTTCTTACCTACCTGGCTCTCATACAGACCCAGAGTGGGGCC 861
QY 263 ProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrpAlaTyr 282
DB 862 CCAGCCAGAGGTCTGGCACTCCAGGCCCTTCAACAGAGAGGAGGATGTGGCTTAC 921
QY 283 TyrGluileCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSerGlnGlu 302
DB 922 TATGAAGTCTGCTCTGG-----AAGGGGGCCACCACACAGAGATCCAGATCAGAAG 975
QY 303 ValProTyrAlaTyrLysAlaAsnGluTrpLeuGlyTyrAspAsnIleLysSerPheSer 322
DB 976 GTGCCCTATCTTCGGGACAAACCAAGTGGTGGGCTTTGTATGATGTGGAGAGCTTCAAA 1035

323 VallysAlaGlnTrpLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTrpAlaIleAsp 342
1036 ACCAAGGTGAGTATCTGAAGCAGAGAGGACTGGGCGGGGCCATGGTCTGGGCACTGGAC 1095
QY 343 LeuAspAspPheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSerThrLeu 362
DB 1096 TTAGATGACTTTGGCGGCTTCTCTGCAACAGGGCCGATACCCCTCATCCAGACGCTA 1155
QY 363 AsnLysAlaLeuGlyIleSerThrGluGlyCysThrAlaProAspVal----- 378
DB 1156 CGGACGAGAACTGAGTCTTCATACCTTCCAGCCACCCAGAGCTTGAAGTTCACAAA 1215
QY 379 -----ProSerGluProValThrThrProGlySerGlySerGlyGlyGlySer 395
DB 1216 CCAGGTGAGCCCTCTGAACCT-----GAG 1239
QY 396 SerGlyGlySerSerGlyGlySerGlyPheCysAlaAspLysAlaAspGlyLeuTyrPro 415
DB 1240 CATGCCCCAGCCCTTGGACACACAGACAGCTTCTGCCAGGCAAGCTGATGGCTCTATCCC 1299
QY 416 ValAlaAspAspArgAsnAlaPheTrpGlnCysIleAsnGlyIleThrTyrGlnGlnHis 435
DB 1300 AATCTCCGGAACGGTCCAGCTTTCAGAGCTGTGCGAGCGGGCGGCTGTCAGCAAGAC 1359
QY 436 CysGlnAlaGlyLeuValPheAspThrSerCysAsnCysCysAsnTyr 451
DB 1360 TGCCCGACAGGCGTGTGTTCAGCAACTCTCTGCAAAATGCTGCACCTGG 1407

RESULT 5
US-09-151-011-3
; Sequence 3, Application US/09151011
; Patent No. 6057142
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: A Human Chitinase, Its Recombinant
; TITLE OF INVENTION: Production, Its Use For Decomposing Chitin, Its Use in
; TITLE OF INVENTION: Therapy or Prophylaxis Against Infection Diseases.
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron, LLP
; STREET: 6900 Jericho Turnpike
; CITY: Syosset
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11791
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/151,011
; FILING DATE: 10 - September - 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Morris, Robert C.
; REGISTRATION NUMBER: 42,910
; REFERENCE/DOCKET NUMBER: 294-32 DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1643 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
US-09-151-011-3
Alignment Scores: 6.74e-121 Length: 1643
Pred. No.:

Score: 1301.50 Matches: 235
Percent Similarity: 67.11% Conservative: 71
Best Local Similarity: 51.54% Mismatches: 129
Query Match: 52.21% Indels: 21
DB: 3 Gaps: 3

US-10-004-219B-9 (1-452) x US-09-151-011-3 (1-1643)

QY 3 LeuileCysThrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPheLysPro 22
DB 82 CTGGTCTGCTACTTCAACCACTGGGCCAGTACAGACAGGGAGGCTCGCTTCCTGGCCC 141
QY 23 AspAspIleAsnProCysLeuCysThrHisLeuileTyrAlaPheAlaGlyMetGlnAsn 42
DB 142 AAGGACTTGGACCCAGCCTTTGGACCCACCTCATCTAGCCCTTCGCTGGCATGCCAAC 201
QY 43 AsnGluileThrThrIleGluTrpAsnAspValThrLeuTyrLysAlaPheAsnAspLeu 62
DB 202 CACCAGCTGAGCACCACCTGAGTGGATGACGAGACTCTCTACAGGAGTTCAATGGCCTG 261
QY 63 LysAsnArgAsnSerLysLeuLysThrLeuLeuAlaileGlyClyTrpAsnPheGlyThr 82
DB 262 AAGAAGATGAATCCCAAGCTTGAAGACCTCTTAGCCATCGAGGCTGGAAATTTGGCACT 321
QY 83 AlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheIleThrSerVal 102
DB 322 CAGAGTTTCAAGATATGTTAGCCAGCGCCCAACACCGTCAGACCTTTCTCACTCGGCC 381
QY 103 IleLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyrProGly 122
DB 382 ATCAGGTTTCTCGCAATACAGCTTTGACGGCTTGTACCTTGTGAGTGGAGTACCAGGA 441
QY 123 SerArgGlySerProGlnAspLysHisLeuPheThrValLeuValLysGluMetArg 142
DB 442 AGCGGGAGGAGCCCTCGCTAGCAAGAGCGCTTCAACACCTGTGTAGCAGACTTGGCC 501
QY 143 GluAlaPheGluGlnGluAlaileGluSerAsnArgProArgLeuMetValThrAlaAla 162
DB 502 AATGCTTTCAGCAGGAGAGCCAGACCTCAGGAGAGGAGCGCTTCTTCTGAGTGCAGCG 561
QY 163 ValAlaGlyGlyIleSerAsnIleGlnAlaGlyTyrGluIleProGluLeuSerLysTyr 182
DB 562 GTTCCAGCTGGGAGAGCTATGTGATGTGTGATAGTGTGAGTGTGAGTGTGAGTGTGAG 621
QY 183 LeuAspPheIleHisValMetThrThrArgLeuLysGlySerTrpGluGlyTyrThrGly 202
DB 622 CTGGATTTTGTCAACCTTATGCGCTACGACTTCCATGGCTCTTGGAGAGGTACGGGA 681
QY 203 GluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeuAsnVal 222
DB 682 CATACAGACCCCTCTACAGAGGCAAGAGAGAGTGTGACAGCAGCAGCTCAACGTG 741
QY 223 AspTyrValMetAsnTyrTrpLysAsnAsnGlyAlaProAlaGluLysLeuIleValGly 242
DB 742 GATGCTGTGTGCAACAGTGGCTGCAAGAGGGGAGCCCTGCCAGCAAGCTGATCTCTTGC 801
QY 243 PheProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIleGlyAla 262
DB 802 ATGCTTACCTAGGAGGCTCTTCACTGGCTCTCTACAGACACCAAGAGTGGGGGCC 861
QY 263 ProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrpAlaTyr 282
DB 862 CCAGCCACAGGCTTGGCACTCCAGGCCCTTCAACCAAGAGGAGGATGCTGGCTAC 921
QY 283 TyrGluileCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSerGlnGlu 302
DB 922 TATGAAGTGTCTCTCTGG-----AAGGGGGGCCCAACAGAGAATCCAGGATCAGAAG 975
QY 303 ValProTyrAlaTyrLysAlaAsnGluTrpLeuGlyTyrAspAsnIleLysSerPheSer 322
DB 976 GTGCCCTCATCTCCGGGACACCAAGCTGGGGCTTTGATGATGGAGAGCTTCAAA 1035
QY 323 ValLysAlaGlnTrpLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTrpAlaIleAsp 342

DB 1036 ACCAAGTTCAGCTATCTGAAGCAGAAGGAGTGGGGCGGGCCATGGTCTGGGCACTGGAC 1095
QY 343 LeuAspAspPheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSerThrLeu 362
DB 1096 TTAGATGACTTTGGCGGCTTCTCTGCAACAGGGCCGATACCCCTCATCCAGAGCGTA 1155
QY 363 AsnLysAlaLeuGlyIleSerThrGluGlyCysThrAlaProAspVal----- 378
DB 1156 CGGACGAGCACTGAGTCTTCATCTTGCCTTCAGCACCCAGAGCTTGAAGTTCCAAA 1215
QY 379 -----ProSerGluProValThrThrProProGlySerGlyGlySer 395
DB 1216 CCAGGTCAGCCCTCTGAACCT-----GAG 1239
QY 396 SerGlyGlySerSerGlySerGlyPheCysAlaAspLysAlaAspGlyLeuTyrPro 415
DB 1240 CATGCCCCAGCCCTTGGACACACAGCTTCTGCCAGGCAAGCTGATGGGCTCTATCCC 1299
QY 416 ValAlaAspAspArgAsnAlaPheTrpGlnCysIleAsnGlyIleThrTyrGlnGlnHis 435
DB 1300 AATCTCCGGACGGTCCAGCTTCTACAGCTGTGACGGGGGGCGCTGTCAGCAAGAC 1359
QY 436 CysGlnAlaGlyLeuValPheAspThrSerCysAsnCysCysAsnTyr 451
DB 1360 TGCCCGACAGCGCTGGTGTTCAGCAACTCTCTGCAAAATGCTGCACCTGG 1407

RESULT 6
US-09-343-623-3
; Sequence 3, Application US/09343623
; Patent No. 6303118
; GENERAL INFORMATION:
; APPLICANT: A human chitinase, its recombinant
; TITLE OF INVENTION: production, its use for decomposing chitin, its use
; TITLE OF INVENTION: in therapy or prophylaxis against infection diseases.
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11758
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/343,623
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,839
; FILING DATE: 07-June-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Baron, Ronald J.
; REGISTRATION NUMBER: 29,281
; REFERENCE/DOCKET NUMBER: 294-26
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1643 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; US-09-343-623-3

Alignment Scores:

Pred. No.: 6,746-121 Length: 1643
Score: 1301.50 Matches: 235
Percent Similarity: 67.11% Conservative: 71
Best Local Similarity: 51.54% Mismatches: 129
Query Match: 52.21% Indels: 21
DB: 4 Gaps: 3

US-10-004-219B-9 (1-452) x US-09-343-623-3 (1-1643)

```
QY 3 LeuileCysThrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPheLysPro 22
DB 82 CTGGTCTCTACTTCCACCACTGGGCCAGTACAGACAGGGGGAGGCTCGCTTCGCGCC 141
QY 23 AspAspIleAsnProCysLeuCysThrHisLeuileTyrAlaPheAlaGlyMetGlnAsn 42
DB 142 AAGGACTTGGACCCAGCCTTGGACCCACCTCATCTAGCCTTCGCTGGATGACCAAC 201
QY 43 AsnGluileThrIleGluTrpAsnAspValThrLeuTyrLysAlaPheAsnAspLeu 62
DB 202 CACCAGCTGAGCACCCTGAGTGGGAATCAGCAGACTCTTACCAGGAGTTCATGGCCCTG 261
QY 63 LysAsnArgAsnSerLysLeuLysThrLeuLeuAlaIleGlyLysTrpAsnPheGlyThr 82
DB 262 AAGAAGATGAATCCCAAGCTGAAGACCTGTGTAGCCATCGGAGGCTGGAATTTGCGCACT 321
QY 83 AlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheIleThrSerVal 102
DB 322 CAGAAGTTCACAGATATGTTAGTCCAGCGCCCAACACCGTCAGACCTTTGTCAACTCGGCC 381
QY 103 IleLysPheLeuArgGlnTrpGlyPheAspGlyLeuAspLeuAspTrpGluTyrProGly 122
DB 382 ATCAGGTTTCTGCGCAATACAGCTTTGACGCGCTTGACCTTGACTGGGAGTACCCAGGA 441
QY 123 SerArgGlySerProGlnAspLysHisLeuPheThrValLeuValLysGluMetArg 142
DB 442 AGCCAGGGAGCCCTGCGCTAGCAGAGGAGCGCTTTCACAAACCCCTGGTACAGGCTGCC 501
QY 143 GluAlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThrAlaAla 162
DB 502 AATGCTCTCCAGCAGAGCCAGCCAGCTCAGGAAGGAGCGCTTCTGTGAGTGCAGCG 561
QY 163 ValAlaGlyGlyIleSerAsnIleGlnAlaGlyTyrGluileProGluLeuSerLysTyr 182
DB 562 GTTCCAGCTGGGCAGACCTATGTGGATGTGTGATACGAGGTGGACAAATGCCCCAGAG 621
QY 183 LeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTrpGluClyTyrThrGly 202
DB 622 CTGGATTTTGTCAACCTTATGGCCTACGACTTCCATGGCTCTTGGGAGAAGGTCAAGGGA 681
QY 203 GluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeuAsnVal 222
DB 682 CATAACAGCCCTCTACAGAGCGCAGAGAGAGTGGTGACAGCCAGCGCTCAACGTG 741
QY 223 AspTyrValMetAsnTyrTrpLysAsnAsnGlyValAlaProAlaGluLysLeuileValGly 242
DB 742 GATGCTGTGTGCAACAGTGGCTGCGAGAAGGGAGCCCTGCCAGCAGACTGATCTTGGC 801
QY 243 PheProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIleGlyAla 262
DB 802 ATGCTTACCTACGAGCGCTCTTCACTTCCACTGGCCTCTCATCAGACACAGAGTGGGGCC 861
QY 263 ProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrpAlaTyr 282
DB 862 CCAGCCAGAGGCTTGGGACTCCAGCGCCCTTCCACCAAGAGAGGAGGATGTCGCCCTAC 921
QY 283 TyrGluileCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSerGlnGlu 302
DB 922 TATGAAGTCTGCTCTGG-----AAGGGGGCCACCAACAGAGATCCAGGATCAGAG 975
QY 303 ValProTyrAlaTyrLysAlaAsnGluTrpLeuGlyTyrAspAsnIleLysSerPheSer 322
DB 976 GTGCCCTACATCTCCGGGACCAACAGTGGGTGGGCTTTGATGATGTGAGAGCTTCAAA 1035
```

```
QY 323 ValLysAlaGlnTrpLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTrpAlaIleAsp 342
DB 1036 ACCAAGTGCAGCTATCTGACAGCAAGGACTGGCGGGGCCATGGTCTGGGCACTGGAC 1095
QY 343 LeuAspAspPheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSerThrLeu 362
DB 1096 TTAGATGACTTTGGCGGCTTCTCCTGCAACACAGGGCGGATACCCGCTCATCCAGACGCTA 1155
QY 363 AsnLysAlaLeuGlyIleSerThrGluGlyCysThrAlaProAspVal----- 378
DB 1156 CGGCAGGAACTGACTTCCATACTTGCCTTCAGGACCCCGAGGCTTGAAGTTCCAAAA 1215
QY 379 -----ProSerGluProValThrThrProGlySerGlySerGlyGlySer 395
DB 1216 CCAGGTGAGCCCTCGAACCT-----GAG 1239
QY 396 SerGlyGlySerSerGlySerGlyPheCysAlaAspLysAlaAspGlyLeuTyrPro 415
DB 1240 CATGGCCCCAGCCCTGGACAGACACGTTCTGCCAGGCAAGCTGATGGGCTCTATCCC 1299
QY 416 ValAlaAspAspArgAsnAlaPheTrpGlnCysIleAsnGlyIleThrTyrGlnGlnHis 435
DB 1300 AATCTCTGGGAACGCTCCAGCTTCTACAGCTGTGACGGGGCGGCTTCCAGCAAGC 1359
QY 436 CysGlnAlaGlyLeuValPheAspThrSerCysAsnCysCysAsnTyr 451
DB 1360 TGCCCGCAGCGGCTGGTTCAGCAACTCTCTGCAAAATGCTGCACCTGG 1407

RESULT 7
US-09-039-198A-3
; Sequence 3, Application US/09039198A
; Patent No. 6200951
; GENERAL INFORMATION:
; APPLICANT: Gray, Patrick W.
; APPLICANT: Tjoelker, Larry W.
; TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive/6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/039,198A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 27866/34391
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1656 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 27..1424
; FEATURE:
; NAME/KEY: mat_peptide
```

LOCATION: 90...1424
US-09-039-198A-3

Alignment Scores:

Pred. No.: 2,72e-120 Length: 1656
Score: 1295.50 Matches: 234
Percent Similarity: 66.89% Conservative: 71
Best Local Similarity: 51.32% Mismatches: 130
Query Match: 51.97% Indels: 21
DB: 4 Gaps: 3

US-10-004-219B-9 (1-452) x US-09-039-198A-3 (1-1656)

```
QY 3 LeuileCysTyrPheThrAsnTrpAlaGlnTyrAtqProGlyLeuGlySerPheLysPro 22
DB 96 CTGGTCTGTACTTCAACCACTGGGCGCCAGTACAGACAGGGGAGGCTCGCTTCCTGCC 155
QY 23 AspAspIleAsnProCysLeuCysThrHisLeuileTyrAlaPheAlaGlyMetGlnAsn 42
DB 156 AAGGACTTGGACCCAGCCTTTGACCCACCTCATCTACGCTTCGCTGGCATGACCAAC 215
QY 43 AsnGluileThrThrileGluTrpAsnAspValThrLeuTyrLysAlaPheAsnAspLeu 62
DB 216 CACCAGCTGAGCACCACTGAGTGAATGACGACTCTTACCGAGGAGTTCAATGGCCGTG 275
QY 63 LysAsnArgAsnSerLysLeuLysThrLeuLeuAlaileGlyLysTrpAsnPheGlyThr 82
DB 276 AGAAGATGATCCCAAGCTGAGACCTTTAGCCATCGAGGCTGGAATTCAGCACT 335
QY 83 AlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheileThrSerVal 102
DB 336 CAGAGTTTCCACAGATATGTTAGCCACCGGCCAACACCGTCAGACCTTTGTCAACTCGGCC 395
QY 103 IleLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyrProGly 122
DB 396 ATCAGGTTTCTGGCAATATACAGCTTTGACGCTTGTACCTTGTACCTGGAGTACCCAGGA 455
QY 123 SerArgGlySerProProGlnAspLysHisLeuPheThrValLeuLysGluMetArg 142
DB 456 AGCCAGGGAGGCGCTCCGTAGACAGGAGCGCTTCAACACCTGGTACAGACTGGCC 515
QY 143 GluAlaPheGluGlnGluAlaileGluSerAsnArgProArgLeuMetValThrAlaAla 162
DB 516 AATGCTTCCAGCAGGAGCCAGACCTCAGGGAAGAACGCTTCTCTGAGTGCAGCG 575
QY 163 ValAlaGlyLysSerAsnIleGlnAlaGlyTyrGluileProGluLeuSerLysTyr 182
DB 576 GTTCCAGCTGGCGACACCTATGTGGATGTGTGATACGAGTGGACAAATCGCCAGAAC 635
QY 183 LeuAspPheileHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThrGly 202
DB 636 CTGGATTTGTCAACCTTATGGCTACGACTCCATGGCTCTTGGGAAGGTACGGGA 695
QY 203 GluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeuAsnVal 222
DB 696 CATAACAGCCCTCTTACAAGAGGCAAGAGAGTGGTGACAGCCAGCCTCAACGTG 755
QY 223 AspTyrValMetAsnTyrTrpLysAsnAsnGlyAlaProAlaGluLysLeuileValGly 242
DB 756 GATGCTGTGTCAACAGTGGCTGCAAGAGGGGACCCCTGCGCAGCAAGCTGATCTTGGC 815
QY 243 PheProGluTyrGlyHisThrPheileLeuArgAsnProSerAspAsnGlyileGlyAla 262
DB 816 ATGCTTACTACGAGCGCTCTTCACTTGGCTCTCTCATCAGACACAGAGTGGGGCC 875
QY 263 ProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrpAlaTyr 282
DB 876 CCAGCCACAGGCTTGGCACTCCAGGCCCTTCAACCAAGAGGAGGATGCTGGCTAC 935
QY 283 TyrGluileCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSerGlnGlu 302
DB 936 TATGAAGTCTGTCTCTGG-----AAGGGGGGCCCAACACAGAGAATCCAGGATCAGAG 989
```

```
QY 303 ValProTyrAlaTyrLysAlaAsnGluTrpLeuGlyTyrAspAsnIleLysSerPheSer 322
DB 990 GTGGCTTACATCTCCGGGACAAACAGTGGGTGGCTTGTATGATGGAGAGCTTCAA 1049
QY 323 ValLysAlaGlnTrpLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTrpAlaileAsp 342
DB 1050 ACCAAGGTGAGTATCTGAAGCAGAGGAGTGGGCGGGGCGCATGCTTGGGCACTGGAC 1109
QY 343 LeuAspAspPheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSerThrLeu 362
DB 1110 TTAGATGACTTTGGCGGCTTCTCTGCAACAGGAGGCGCATACCCCTCATCCAGACGCTA 1169
QY 363 AsnLysAlaLeuGlyIleSerThrGluGlyCysThrAlaProAspVal----- 378
DB 1170 CGGACGAGTACGAGTCTTCATACCTTGCCTTCAGCACCCCGAGAGCTTGAAGTCCAAAA 1229
QY 379 -----ProSerGluProValThrThrProGlySerGlyGlySer 395
DB 1230 CCAGGTCAGCCCTCTGAACCT-----GAG 1253
QY 396 SerGlyGlySerSerGlyGlySerGlyPheCysAlaAspLysAlaAspGlyLeuTyrPro 415
DB 1254 CATGCCCGCAGCCCTGGACACACAGCTTCTGCCAGGCGCAAGCTGATGGGCTCTATCCC 1313
QY 416 ValAlaAspAspArgAsnAlaPheTrpGlnCysIleAsnGlyIleThrTyrGlnGlnHis 435
DB 1314 AATCTTCGGGAACGCTCCAGCTTCTACAGCTGTGCAGCGGGCGGCTGTTCCAGCAAGC 1373
QY 436 CysGlnAlaGlyLeuValPheAspThrSerCysAsnCysCysAsnTrp 451
DB 1374 TGCCCGCAGCAGCGCTGGTGTTCAGCAACTCTCTGCAAAATGCTGCACCTGG 1421
```

RESULT 8

US-08-877-599-3
Sequence 3, Application US/08877599
Patent No. 6372212
GENERAL INFORMATION:
APPLICANT: Gray, Patrick W.
TITLE OF INVENTION: Chitinase Materials and Methods
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/877,599
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/663,618
FILING DATE: 14-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Rin-Laures, Li-Hsien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 27866/33994
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1656 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 27..1424
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 90..1424
US-08-877-599-3

Alignment Scores:

Pred. No.: 2,72e-120 Length: 1656
Score: 1295.50 Matches: 234
Percent Similarity: 66.89% Conservative: 71
Best Local Similarity: 51.32% Mismatches: 130
Query Match: 51.97% Indels: 21
DB: 4 Gaps: 3

US-10-004-219B-9 (1-452) x US-08-877-599-3 (1-1656)

QY 3 LeuileCysThrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPheLysPro 22
DB 96 CTGGTCTGCTACTTCCCACTGGCCCGAGTACAGACAGGGGGAGGCTCGCTTCTCGGCC 155
QY 23 AspAlaPheAsnProCysLeuCysThrHisLeuileTyrAlaPheAlaGlyMetGlnAsn 42
DB 156 AAGGACTTGGACCCAGCCCTTTCACCCACCTCATCTAGCCCTTCGCTGGCATGCCAAC 215
QY 43 AsnGluileThrIleGlnTrpAsnValThrLeuTyrLysAlaPheAsnAspLeu 62
DB 216 CACCGCTGAGCACCAGTGGTGAATGACGAGCTCTTACCGAGGTTCAATGGCCGTG 275
QY 63 LysAsnArgAsnSerLysLeuLysThrLeuLeuAlaIleGlyLysTrpAsnPheGlyThr 82
DB 276 AAGAGATCAATCCCACTGAGACCTGTAGCCATCGAGGCTGGAAATTCAGCACT 335
QY 83 AlaProPheThrMetValSerThrSerGlnAsnArgGlnThrPheIleThrSerVal 102
DB 336 CAGAAAGTTCACAGATATGTTAGCCAGCGCAACACCGTCAGACCTTTGTCAACTCGGCC 395
QY 103 IleLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGlnTyrProGly 122
DB 396 ATCAGGTTTCTCGCAATACAGCTTTGACGCCCTTGACCTTGAGTGGAGTACCCAGGA 455
QY 123 SerArgGlySerProGlnAspLysHisLeuPheThrValLeuValLysGluMetArg 142
DB 456 AGCCAGGGAGGAGCCCTCCCTAGACAGAGGCGCTTTCACACCTCGTACAGACTTGCC 515
QY 143 GluAlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThrAlaAla 162
DB 516 AATGCTTCCAGCAGGAGCCAGACCTCAGGGAAGGAACGGCTTCTTCTGAGTGCAGCG 575
QY 163 ValAlaGlyGlyLysSerAsnIleGlnAlaGlyTyrGluileProGluLeuSerLysTyr 182
DB 576 GTTCCAGCTGGCGACACCTATGTTGATGATGATGATGATGATGATGATGATGATGATG 635
QY 183 LeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThrGly 202
DB 636 CTGGATTTTGTCAACCTTAGCCCTACGACTTCCATGGCTCTTGGAGAGGTACGGGA 695
QY 203 GluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeuAsnVal 222
DB 696 CATAACAGCCCTCTTACAAGAGGCAAGAGAGTGGTGCAGCAGCCAGCCCTCAACGTG 755
QY 223 AspTyrValMetAsnTyrTrpLysAsnAsnGlyAlaProAlaGluLysLeuileValGly 242
DB 756 GATGCTGTGTCAACAGTGGCTGTCAGAGGGGAGCCCTCGCCAGCAAGCTGATCTTGGC 815
QY 243 PheProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIleGlyAla 262
DB 816 ATGCTTACTACGAGCGCTTCTTACACTGGCTCTCTATCAGACACCCAGAGTGGGGCC 875
QY 263 ProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrpAlaTyr 282

DB 876 CCAGCCACAGGGTCTGGCACTCCAGGCCCTTCCACAGGAAGGAGGATGCTGGCCTAC 935
QY 283 TyrGluileCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSerGlnGlu 302
DB 936 TATGAAGTCTGCTCTCTGG-----AAGGGGCCACCAACACAGAAATCCAGATCAGAAG 989
QY 303 ValProTyrAlaTyrLysAlaAsnGlnTrpLeuGlyTyrAspAsnIleLysSerPheSer 322
DB 990 GTGCCCTACATCTTCGGGACACACAGTGGGTGGCTTTGATGATGTGGAGAGCTTCAAA 1049
QY 323 ValLysAlaGlnTrpLeuLysGlnAsnAsnPheGlyAlaMetIleTrpAlaIleAsp 342
DB 1050 ACCAAGGTCACTATCTGAAGCAGAGGGGACTGGGGCGGGCCATGGTCTGGGCACTGGAC 1109
QY 343 LeuAspAspPheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSerThrLeu 362
DB 1110 TTAGATGATTTTGGCGGCTTCTCTGCAACAGGGGCGCATACCCCTCATCCAGAGGCTA 1169
QY 363 AsnLysAlaLeuGlyIleSerThrGluGlyCysThrAlaProAspVal----- 378
DB 1170 CGGACGGAACCTGAGTCTTCCATACTTGGCTTTCAGGCACCCAGAGCTTGAAGTTCAAAA 1229
QY 379 -----ProSerGluProValThrThrProGlySerGlySerGlyGlySer 395
DB 1230 CCAGTCTCAGCCCTCTGACCT-----GAG 1253
QY 396 SerGlyGlySerSerGlySerGlyPheCysAlaAspLysAlaAspGlyLeuTyrPro 415
DB 1254 CATGCCCCACCCCTGGACAAAGACACAGCTTCTGCCAGGCAAGCTGATGGCTCTATCCC 1313
QY 416 ValAlaAspAspArgAsnAlaPheTrpGlnCysIleAsnGlyIleThrTyrGlnGlnHis 435
DB 1314 AATCTCGGGAACGGTCCAGCTTCTACAGCTGTGCGGGGGCGGCTGTTCAGCAAGAC 1373
QY 436 CysGlnAlaGlyLeuValPheAspThrSerCysAsnCysCysAsnTrp 451
DB 1374 TGCCGACAGGCGCTGGTGTTCAGCAACTCTCTGCAAAATGCTGCACCTGG 1421

RESULT 9

US-09-267-574-3
; Sequence 3, Application US/09267574
; Patent No. 6399571
; GENERAL INFORMATION:
; APPLICANT: Gray, Patrick W.
; APPLICANT: Tjoelker, Larry W.
; TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
; FILE REFERENCE: 27866/35407
; CURRENT APPLICATION NUMBER: US/09/267,574
; CURRENT FILING DATE: 1999-03-12
; EARLIER APPLICATION NUMBER: 09/039,198
; EARLIER FILING DATE: 1998-03-12
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1656
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (27)..(1424)
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (90)..(1424)
US-09-267-574-3

Alignment Scores:

Pred. No.: 2,72e-120 Length: 1656
Score: 1295.50 Matches: 234
Percent Similarity: 66.89% Conservative: 71
Best Local Similarity: 51.32% Mismatches: 130
Query Match: 51.97% Indels: 21
DB: 4 Gaps: 3

QY	3	LeuIleCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPheLysPro	22
DB	82	CTGTCCTGCTACTTCAACCACTGGCGCCAGTACAGACAGGGGAGGCTGCTTCTCTGCGCC	141
QY	23	AspAspIleAsnProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyMetGlnAsn	42
DB	142	AAGAGCTTGGACCCACAGCTTTGACCCACCTCATCTACGCTTCTGCTGCATGACCAAC	201
QY	43	AsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrLysAlaPheAsnAspLeu	62
DB	202	CACCAGCTGAGCACCACTGAGTGGATGACGACACTCTACACAGGAGTTCAATGCGCTG	261
QY	63	LysAsnArgAsnSerLysLeuLysThrLeuLeuAlaIleGlyTrpAsnPheGlyThr	82
DB	262	AAGAAGATGAATCCCAAGCTGAAGACCTGTAGCCATCGAGGCTGGAATTTGCGCACT	321
QY	83	AlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheIleThrSerVal	102
DB	322	CAGAAGTTACAGATATGGTAGCCACCGCCCAACACCGTCAGACCTTGTGCACTCGGCC	381
QY	103	IleLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyrProGly	122
DB	382	ATCAGGTTTCTGGCAATACACTTTGACGGCTTGACCTTGACTGGGAGTACCCAGGA	441
QY	123	SerArgLysSerProGlnAspLysHisLeuPheThrValLeuValLysGluMetArg	142
DB	442	AGCCAGGGAGCCCTGCGGTAGACAAGGAGCGCTTCACAACCTGTGTACAGGACTTGGCC	501
QY	143	GluAlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThrAlaAla	162
DB	502	AATGCTTTCAGCAGGAAGCCACAGACTCCTCAGGGAAGAACGCCCTTCTCTGAGTCAGCG	561
QY	163	ValAlaGlyIleSerAsnIleGlnAlaGlyTyrGluIleProGluLeuSerLysTyr	182
DB	562	GTTCAGCTGGGAGACCTATGTGGATGCTGGATACAGGTGAGCAAAATCGCCCAAGAAC	621
QY	183	LeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThrGly	202
DB	622	CTGGATTTTGTCAACCTTATGGCTACGACTTCATGGCTCTTGGGAGAAGGTACCGGA	681
QY	203	GluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeuAsnVal	222
DB	682	CATAACAGCCCCCTCTCAAGAGAGCAAGAGAGAGTGTGTGTCAGCAGCCCTCAACGTG	741
QY	223	AspTyrValMetAsnTyrTrpLysAsnAsnGlyAlaProAlaGluLysLeuIleValGly	242
DB	742	GATGCTGCTGTCAACAGTGGCTGCAGAGGGGACCCCTCCAGCAAGCTGATCTTGGC	801
QY	243	PheProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIleGlyAla	262
DB	802	ATGCCTACCTACGGAGCGCTCTTCACACTGGCTCTCATCAGACACAGAGTGGGGGCC	861
QY	263	ProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrpAlaTyr	282
DB	862	CCAGCCACAGGCTGTGGCACTCCAGGCCCTTCCAAAGGAAGAGGATGTCTGGCTAC	921
QY	283	TyrGluIleCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSerGlnGlu	302
DB	922	TATGAAGTGTGCTCCTGG-----AAGGGGGGCCACCAACAGAGAATCCAGGATCAAG	975
QY	303	ValProTyrAlaTyrLysAlaAsnGluTrpLeuGlyTyrAspAsnIleLysSerPheSer	322
DB	976	GTGGCTTACATCTTCGGGGCAACCACTGGGTGGCTTGTATGATGTGGAGAGCTTCAA	1035
QY	323	ValLysAlaGlnTrpLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTrpAlaIleAsp	342
DB	1036	ACCAAGTTCAGCTATCTGAAGCAGAGAGGAGCTGGCGGGGCCATGTGCTGGGCACTGG	1095
QY	343	LeuAspAspPheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSerThrLeu	362
DB	1096	TTAGATGACTTTGCCGCTTCTCTCTGCAACAGGGCGGATACCCCTCTCATCCAGACGCTA	1155

QY	363	AsnLysAlaLeu-----	-----GlyIleSer	369
		:::	::::	
Db	1156	CGCGAGCAACTGAATGGGTAAAGCCCTTA	ACTGCTCTCACATGTGAGGCCAGGTGTGCC	1215
QY	370	ThrGluGlyCys---ThrAlaProAspValProSerGluProValThrThrPro-----	-----	386
Db	1216	TGTGG-CACTGTGCTTCACTGTAGTCTCTCCATATCTGCCTTCAGGCACCCAGAGCTT	1274	
QY	387	-----ProGlySerGlySerGlyGlySerSerGlyGlySerSerGlyGly 402		
Db	1275	GAAGTTCCAAACACCGAGTTCAGCCCTCT--GAACCTGAGCATGGCCCAAGCCCTGGACAA	1331	
QY	403	SerGlyPheCysAlaAspLysAlaAspGlyLeuTyProValAlaAspAspArgAsnAla	422	
Db	1332	GACACGCTTCTGCCAGGCGCAAGCTGATGGGCTCTATCCCAATCTCGGGAAACGGTCCAGC	1391	
QY	423	PheTrpGlnCysIleAsnGlyIleThrTyTrGlnGlnHisCysGlnAlaGlyLeuValPhe	442	
Db	1392	TTCTACAGCTGTGCACGGGGGGGGCTGTTCACGAAAGCTGCCGACAGCCCTGGTGTTC	1451	
QY	443	AspThrSerCysAsnCysCysAsnTrp	451	
Db	1452	AGCAACTCTCGCAATGTGTCACCTGG	1478	

RESULT 11
 US-09-151-011-5
 ; Sequence 5, Application US/09151011
 ; Patent No. 6057142
 ; GENERAL INFORMATION:
 ; APPLICANT:
 ; TITLE OF INVENTION: A Human Chitinase, Its Recombinant
 ; TITLE OF INVENTION: Production, Its Use For Decomposing Chitin, Its Use i
 ; TITLE OF INVENTION: Therapy or Prophylaxis Against Infection Diseases.
 ; NUMBER OF SEQUENCES: 16
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Hoffmann & Baron, LLP
 ; STREET: 6900 Jericho Turnpike
 ; CITY: Syosset
 ; STATE: New York
 ; COUNTRY: United States of America
 ; ZIP: 11791
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; SOFTWARE:
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/151,011
 ; FILING DATE: 10 - September - 1998
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Morris, Robert C.
 ; REGISTRATION NUMBER: 42,910
 ; REFERENCE/DOCKET NUMBER: 294-32 DIV
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (516) 822-3550
 ; TELEFAX: (516) 822-3582
 ; INFORMATION FOR SEQ ID NO: 5:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1713 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: unknown
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: cDNA
 ; HYPOTHETICAL: NO
 ; US-09-151-011-5

Alignment Scores:			
Pred. No.:	2,57e-119	Length:	1713
Score:	1286.00	Matches:	240
Percent Similarity:	66.52%	Conservative:	72
Best Local Similarity:	51.17%	Mismatches:	134
Query Match:	51.58%	Indels:	24

DB: 3 Gaps: 5
US-10-004-219B-9 (1-452) x US-09-151-011-5 (1-1713)
QY 3 LeuileCysThrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPheLysPro 22
DB 82 CTGGTCTGTACTTCAACAACTGGGCGCCAGTACAGACAGGGGAGAGCTCGCTTCTCTGCC 141
QY 23 AspAspIleAsnProCysLeuCysThrHlaLeuLeuTyrAlaPheAlaGlyMetGlnAsn 42
DB 142 AAGGACTTGGACCCAGCCTTTCACCCACCTCATCTAGCCCTTCCTGGCATGACCAAC 201
QY 43 AsnGluileThrThrileGluTrpAsnAspValThrLeuTyrIlysAlaPheAsnAspLeu 62
DB 202 CACCAGCTGAGCACCACCTGAGTGGATGACGAGACCTCTTACCAGGAGTTCAATGSCCTG 261
QY 63 LysAsnArgAsnSerIlysLeuIysThrLeuLeuAlaileGlyTrpAsnPheGlyThr 82
DB 262 AAGAAGATGAATCCCAAGCTGAAGACCTGTTAGCCATCGGAGGCTTGAATTTGGCACT 321
QY 83 AlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheIleThrSerVal 102
DB 322 CAGAACTTACAGATATGGTAGCCAGCGCCACACACCTGACACCTTTGTCAACTCGGCC 381
QY 103 IleLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyrProGly 122
DB 382 ATCAGGTTTCTGCGCAATACAGCTTTGACGGCCTTGACCTTGAGTGGGAGTACCCAGGA 441
QY 123 SerArgLysSerProGlnAspLysHlaLeuPheThrValLeuValLysGluMetArg 142
DB 442 AGCAGGGAGGAGCCCTCCGCTAGACAGGAGCGCTTCAACACCTCGTACAGACTTTGGCC 501
QY 143 GluAlaPheGluGlnGluAlaileGluSerAsnArgProArgLeuMetValThrAlaAla 162
DB 502 ATGCTCTTCAGCGAAGCCAGACTCAGGAGGAAGACGCTTCTCTGAGTGACGG 561
QY 163 ValAlaGlyIleSerAsnIleGlnAlaGlyTyrGluileProGluLeuSerIysTyr 182
DB 562 GTTCCAGCTGGGAGACCTATGTGGATGTGGATACGAGTGGACAAAATCGCCAGAAC 621
QY 183 LeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThrGly 202
DB 622 CTGGATTTTGTCAACCTTATGGCTACGACTTCCATGGCTCTTGGGAGAGGTACGGGA 681
QY 203 GluAsnSerProLeuTyrIysTyrProThrGluThrGlySerAsnAlaTyrLeuAsnVal 222
DB 682 CATACAGGCCCTCTACAGAGCGCAAGAGAGAGTGGTGACAGCCAGCCTCAACGTG 741
QY 223 AspTyrValMetAsnTyrTrpLysAsnAsnGlyAlaProAlaGluLysLeuileValGly 242
DB 742 GATGCTGTGTGCAACAGTGGCTGCAGAGGGGAGCCCTGCCAGCAAGCTGATCCTTGGC 801
QY 243 PheProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIleGlyAla 262
DB 802 ATGCTCTACCTGAGGAGCTCTTCACTACCTGGCTCTCATACAGACACAGAGTGGGGGCC 861
QY 263 ProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrpAlaTyr 282
DB 862 CCAGCCACAGGCTGTGGACTCTCAGGCCCTTCCACAGAGAGGAGGATGCTGGCTTAC 921
QY 283 TyrGluileCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSerGlnGlu 302
DB 922 TATGAAGTGTCTCTGG-----AAGGGGGCCCAACACAGAGATCCAGGATCAGAG 975
QY 303 ValProTyrAlaTyrIysAlaGlnGluTrpLeuGlyTyrAspAsnIleLysSerPheSer 322
DB 976 GTGCCCTATCTTCCGGGACCAACAGTGGGCTTGTGATGATGGAGAGCTTCAAA 1035
QY 323 ValLysAlaGlnTrpLeuLysGlnAsnAsnPheGlyGlyValaMetIleTrpAlaIleAsp 342
DB 1036 ACCAAGTCAGTATCTGAAGCAGAGAGGACTGGGGGGGGGCTGCTCTGGGACTGGAC 1095
QY 343 LeuAspAspPheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSerThrLeu 362

DB 1096 TTAGATGACTTTGGCGGCTTCTCTGCAACAGGCGCGATACCCCTCATCCAGACGCTA 1155
QY 363 AsnLysAlaLeu-----GlyIleSer 369
DB 1156 CGGCAGGAACCTAATGGTAAAGCCCTTAAGCTTCACATGTGAGGCCAGGTGTTC 1215
QY 370 ThrGluGlyCys---ThrAlaProAspValProSerGluProValThrPro----- 386
DB 1216 TGTGG-CACTGTGCTTACGCTTAGTCTTCCATCTTCCACTTTCAGGACCCAGAGCTT 1274
QY 387 -----ProGlySerGlySerGlyGlySerSerGlyGlySerGlyGly 402
DB 1275 GAAGTTCACAAAACAGGTGAGCCCTCT---GAACCTGAGCATGGCCCAAGCCCTGCACAA 1331
QY 403 SerGlyPheCysAlaAspLysAlaAspGlyLeuTyrProValAlaAspArgAsnAla 422
DB 1332 GACAGCTTCTCCAGGGCAAGCTGATGGCTCTATCCCAATCTCTCGGGAACGGTCCAGC 1391
QY 423 PheTrpGlnCysIleAsnGlyIleThrTyrGlnGlnHisCysGlnAlaGlyLeuValPhe 442
DB 1392 TTCTACAGCTGTGACGGGGGGCTGTTCAGCAAAAGCTGCCGACAGGCTGTGTTC 1451
QY 443 AspThrSerCysAsnCysCysAsnTrp 451
DB 1452 AGCAACTCTGCAAAATGCTGCACCTGG 1478

RESULT 12

US-09-343-623-5
; Sequence 5, Application US/09343623
; Patent No. 6303118
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: A human chitinase, its recombinant
; TITLE OF INVENTION: production, its use for decomposing chitin, its use
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Hoffmann & Baron
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11758

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; SOFTWARE:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/343,623

FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,839

FILING DATE: 07-June-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Baron, Ronald J.

REGISTRATION NUMBER: 29,281
; REFERENCE/DOCKET NUMBER: 294-26
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550

TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1713 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO

US-09-343-623-5

Alignment Scores:

Pred. No.: 2,57e-119 Length: 1713
Score: 1286.00 Matches: 240
Percent Similarity: 66.52% Conservative: 72
Best Local Similarity: 51.17% Mismatches: 134
Query Match: 51.58% Indels: 24
DB: 4 Gaps: 5

US-10-004-219B-9 (1-452) x US-09-343-623-5 (1-1713)

```
OY 3 LeuileCysThrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPheIlePro 22
DB 82 CTGGCTCTCTACTTCCCAACTGGGCCAGTACACAGAGGGAGGCTCGCTTCGCGCC 141
OY 23 AspaPileAsnProCysLeuCysThrHisLeuileTyrAlaPheAlaGlyMetGlnAsn 42
DB 142 AAGGACTTGGACCCAGCCTTTCACCCACCTCATCTAGGCTTCGCTGGCATGCCAAC 201
OY 43 AsnGluileThrIleGluTrpAsnAspValThrLeuTyrLysAlaPheAsnAspLeu 62
DB 202 CACCAGCTGAGCACCACCTGAGTGGAAATGACGAGACTCTTACGAGGAGTTCAATGGCCTG 261
OY 63 LysAsnArgAsnSerLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPheGlyThr 82
DB 262 AAGAAGATGAATCCCAAGCTGAAGACCTGTGTAGCCATCGGAGGCTGGAATTCGGCACT 321
OY 83 AlaProPheThrMetValSerThrSerGlnAsnArgGlnThrPheIleThrSerVal 102
DB 322 CAGAATTCACAGATATGGTAGCCAGGCCAACACCGTTCAGACCTTTGTCACTCGGCC 381
OY 103 IleLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyrProGly 122
DB 382 ATCAGGTTTCTGGCAATACAGCTTTGAGCGCTTGACCTTGTGAGTGGAGTACCCAGGA 441
OY 123 SerArgGlySerProGlnAspLysHisLeuPheThrValLeuValLysGluMetArg 142
DB 442 AGCCAGGGGAGCCCTGCGCTAGACAGGAGCGCTTTCACACCCCTGTGTACAGACTTGCC 501
OY 143 GluAlaPheGluGlnAlaIleGluSerAsnArgProArgLeuMetValThrAlaAla 162
DB 502 AATGCTTTCAGCAGAGAGCCAGACCTCAGGGAAGGAGCGCTTCTCTGAGTGACGG 561
OY 163 ValAlaGlyGlyIleSerAsnIleGlnAlaGlyTyrGluileProGluLeuSerLysTyr 182
DB 562 GTTCCAGCTGGGCAGACCTATGTGATGTGTGATGAGGTGACAGCAAAATCCGCCAGAC 621
OY 183 LeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThrGly 202
DB 622 CTGGATTTTGTCAACCTTATGGCCTACGACTTCCATGGCTCTTGGGAGAAGGTCAAGGA 681
OY 203 GluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeuAsnVal 222
DB 682 CATAACAGCCCTCTACAGGCGCAGAGAGAGTGTGTGAGCAGCAGCCTCAACGTG 741
OY 223 AsptyrValMetAsnTyrTrpLysAsnGlnAlaProAlaGluLysLeuileValGly 242
DB 742 GATGCTGTGTCAACAGTGGCTGCAGAGGGGAGCCCTCGCCAGCAAGTGTATCCTTGGC 801
OY 243 PheProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIleGlyAla 262
DB 802 ATGCCCTAGCTACGGAGCGCTCTTCACTACCTGGCCCTCTCATCAGACACCAAGAGTGGGGCC 861
OY 263 ProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrpAlaTyr 282
DB 862 CAGCCACAGGCTCTGGCACTCCAGGCCCTTCCACAGGAAGAGGAGGATGTCGCCCTAC 921
OY 283 TyrGluileCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSerGlnGlu 302
DB 922 TATGAAGTCTGCTCTCGG-----AAGGGGGCCACCAACAGAGATCCAGATCAGAG 975
OY 303 ValProTyrAlaTyrLysAlaAsnGluTrpLeuGlyTyrAspAsnIleLysSerPheSer 322
DB 976 GTGCCCTACATCTCCGGGACCAACCAAGTGGGGGCTTTGATGATGTGGAGAGCTTCAA 1035
```

```
OY 323 VallysAlaGlnTrpLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTrpAlaIleAsp 342
DB 1036 ACCAGGTGACCTATCTCAACAGAGGACTGGGGGGCCATGGTCTGGGCACGTGAC 1095
OY 343 LeuAspAspPheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSerThrLeu 362
DB 1096 TTAGATGACTTTGGCGGCTTCTCTCTGCAACCAAGGCGCGATACCCCTCATCCAGACGCTA 1155
OY 363 AsnLysAlaLeu-----GlyIleSer 369
DB 1156 CGGAGGAACTGAATGGGTAAAGCTTAACTGCTGTGCATGTGAGGCCAGGTGTGCC 1215
OY 370 ThrGluGlyCys---ThrAlaProAspValProSerGluProValThrThrPro----- 386
DB 1216 TGTGG-CACGTGTCTCAGCTGTAGGTCTTCATACCTTTCCTTCCAGGACCCAGAGCTT 1274
OY 387 -----ProGlySerGlySerGlyGlySerGlySerGlySerGlySerGlyGly 402
DB 1275 GAAGTTCCAAACCAAGGTTCAGCCCTCT---GAACCTGAGCATGGCCCAAGCCTCGCAA 1331
OY 403 SerGlyPheCysAlaAspLysAlaAspGlyLeuTyrProValAlaAspAspArgAsnAla 422
DB 1332 GACAGTTCTCCAGGCGCAAGCTGTATGGGCTCTATCCCAATCTCTCGGGAACGGTCCAGC 1391
OY 423 PheTrpGlnCysIleAsnGlyIleThrTyrGlnGlnHisCysGlnAlaGlyLeuValPhe 442
DB 1392 TTCTACAGCTGTGACGGGGGGGGCTGTTCAGCAGCAAGCTGCCCGACAGGCTGTGTGTT 1451
OY 443 AspThrSerCysAsnCysCysAsnTrp 451
DB 1452 AGCAACTCTGCAAAATGCTGCACCTGG 1478
```

RESULT 13

PCT-US94-07754-4

; Sequence 4, Application PC/TUS9407754

; GENERAL INFORMATION:

; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA

; TITLE OF INVENTION: ASSAY FOR YKL-40 AS A MARKER FOR

; TITLE OF INVENTION: DEGRADATION OF MAMMALIAN CONNECTIVE TISSUE MATRICES

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SPENSLEY HORN JUBAS & LUBITZ

; STREET: 1880 CENTURY PARK EAST, FIFTH FLOOR

; CITY: LOS ANGELES

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 90067

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US94/07754

; FILING DATE: 08-JUL-1994

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: HOWELLS, STACY L.

; REGISTRATION NUMBER: 34,842

; REFERENCE/DOCKET NUMBER: FD 3665

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 619/455-5100

; TELEFAX: 619/455-5110

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1681 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; IMMEDIATE SOURCE:

; CLONE: YKL-40

QY 301 GlnGluValProTyrAlaTyrLysAlaAsnGluLutpLeuGlyTyrAspAsnIleLysSer 320

BEST LOCAL SIMILARITY: 43.48%

Alignment Scores:	8.07e-86	Length:	1433
Pred. No.:	950.50	Matches:	181
Score:	66.08%	Conservative:	82
Percent Similarity:	45.48%	Mismatches:	122
Best Local Similarity:			

```
Query Match: 38.13% Indels: 13
DB: 1 Gaps: 6
US-10-004-219B-9 (1-452) x US-08-694-915-1 (1-1433)

QY 1 TyrAsnLeuLeuCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPhe 20
DB 115 TACAAATCGTGTGCTACTTTACCAACTGGTCCAGGACCGGAGGAAACCCAGGAAATTC 174
QY 21 LysProAspAspIleAsnProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet 40
DB 175 ACCCTCGAGAATATTGACCCCTTCCTGCTCATCTCATCTATTCATTCTCGGCAGCATC 234
QY 41 GlnAsnGlnIleThrThrIleGluTrpAsnAspValThrLeuTyrLysAlaPheAsn 60
DB 235 GAAACACACAGGTATCATCAGGACAGAGTGAAGTGTCTTACAGACCATCAAC 294
QY 61 AspLeuLysAsnArgAsnSerLysLeuLysThrLeuLeuAlaIleGlyLysTrpAsnPhe 80
DB 295 AGTCTCAAAACCAAGAAATCCAAACTGAAATTCCTGTGTCATTGGAGGGTACCTGTTT 354
QY 81 GlyThrAlaProPheThrThrMetValSerThrSerClnAsnArgGlnThrPheIleThr 100
DB 355 GGTTCCAAAGGTTCCACCCATGGTGGATTCTTCTACATCACGCTTGGAAATTCATTAAC 414
QY 101 SerValIleLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyr 120
DB 415 TCCATATCTCTGTTCTGAGAACCATTAATTTGATGGACTGGATGTAAGCTGATCTAC 474
QY 121 ProGlySerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValLysGlu 140
DB 475 CCAGATCAGAAA-----GAAACACATCAT---TTCACCTGTGCTGATTATGAG 519
QY 141 MetArgGluAlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThr 160
DB 520 TTAGCAGAGACCTTTTCCAGAGGACTTCCAAATAATCCAAAGGAAAGGCTTCTTTGACT 579
QY 161 AlaAlaValAlaGlyGlyIleSerAsnIleGlnAlaGlyTyrGluIleProGluLeuSer 180
DB 580 GCGGGGTATCTCAGGGAGGCAATGATTTGATAACAGCTATCAAGTTGAGAACTGGCA 639
QY 181 LysTyrLeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTrpGluGly--- 199
DB 640 AAAGATCTGGATTTCATCAACCTCTCTGCTGCTTGGCTTCCATGGGTCTTGGGAAAGCCC 699
QY 200 ---TyrThrGlyGluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAla 218
DB 700 CTTATCACTGGCCACACAGCCCTCTGAGCAGGGGTGGCAGGAGGAGGAGGAGGAGGAGG 759
QY 219 TyrLeuAsnValAspTyrValMetAsnTyrTrpLysAsnAsnGlyAlaProAlaGluLys 238
DB 760 TACTACATGTGGAATATGCTGTGGGGTACTGGATACATAAGGGAATGCCATCAGAGAGAG 819
QY 239 LeuIleValGlyPheProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsn 258
DB 820 GTGGTCATGGGCATCCCAATATGGGCACTCTTCCACATGGGCTCT---GCAGAAACC 876
QY 259 GlyIleGlyAlaProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGly 278
DB 877 ACCGTGGGGGGCCCTGCCTCTGGCCCTGGAGCTGGAGCCCATCAGAGTCTTCAGGC 936
QY 279 PheTrpAlaTyrTyrGluIleCysThrPheLeuArgSerGlyAlaThrGluValTrpAsp 298
DB 937 TTCCTGGCCCTATTATGAGATCTGCCAGTTCCTGAAA---GGAGCCAAAGATCAGCGGCTC 993
QY 299 AlaSerGlnGluValProTyrAlaTyrLysAlaAsnGluTrpLeuGlyTyrAspAsnIle 318
DB 994 CAGGATCAGCAGGTTCCCTACCGAGTCAAGGGGAACAGTGGGTGGGCTATGATGATGTG 1053
QY 319 LysSerPheSerValLysAlaGlnTrpLeuLysGlnAsnAsnPheGlyGlyAlaMetIle 338
DB 1054 AAGATGTGAGAGCAAGGTTGATTTCTTAAGAAATTTTAAACCTGGGAGGAGGAGCCATGATC 1113
```

Search completed: June 30, 2003, 01:11:17
Job time : 95.5237 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 29, 2003, 23:54:39 ; Search time 164.82 Seconds
(without alignments)
4070.941 Million cell updates/sec

Title: US-10-004-219B-9

Perfect score: 2493

Sequence: 1 YNLICYFTWAGYRPLGLSF.....QQHCOAGLVFDTSCNCNWP 452

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1055720 seqs, 742224136 residues

Total number of hits satisfying chosen parameters: 2111440

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=slp
-Q/cgn2_1/USPTO_spool/US10004219/runat_24062003_160230_8285/app.query.fasta_1.1877
-DB=Published Applications NA -QFMT=fastcap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10004219 @CGN 1 1 123 @runat_24062003_160230_8285
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCk=100
-LONGLOG -DEV -TIMEOUT=120 -WARN -TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-XGAPOP=6 -XGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- Published Applications NA:*
- 1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
 - 2: /cgn2_6/prodata/1/pubpna/PCT_NEW_PUB.seq:*
 - 3: /cgn2_6/prodata/1/pubpna/US06_PUBCOMB.seq:*
 - 4: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
 - 5: /cgn2_6/prodata/1/pubpna/PCTUS_PUBCOMB.seq:*
 - 6: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:*
 - 7: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:*
 - 8: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
 - 9: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
 - 10: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
 - 11: /cgn2_6/prodata/1/pubpna/US10_NEW_PUB.seq:*
 - 12: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
 - 13: /cgn2_6/prodata/1/pubpna/US60_NEW_PUB.seq:*
 - 14: /cgn2_6/prodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2493	100.0	1525	9	US-10-004-219B-3
2	2089.5	83.8	1625	9	US-10-004-219B-2
3	992.5	39.8	1925	9	US-10-097-340-44
4	987.5	39.6	1038	9	US-10-202-436A-14

5	966.5	38.8	1681	10	US-09-215-077A-4	Sequence 4, Appli
6	966.5	38.8	1681	10	US-09-262-213A-4	Sequence 4, Appli
7	902.5	36.2	1391	10	US-09-822-830A-402	Sequence 402, App
8	854.5	34.3	1474	10	US-09-765-231A-8	Sequence 8, Appli
9	729.5	29.3	1608	9	US-10-218-743-20	Sequence 20, Appli
c 10	729.5	29.3	1608	9	US-10-218-743-22	Sequence 22, Appli
c 11	729.5	29.3	1665	9	US-10-218-743-17	Sequence 17, Appli
c 12	729.5	29.3	1665	9	US-10-218-743-19	Sequence 19, Appli
c 13	729.5	29.3	1752	9	US-10-218-743-14	Sequence 14, Appli
c 14	729.5	29.3	1752	9	US-10-218-743-16	Sequence 16, Appli
c 15	722	29.0	1470	9	US-10-218-743-40	Sequence 40, Appli
c 16	722	29.0	1470	9	US-10-218-743-42	Sequence 42, Appli
c 17	722	29.0	1527	9	US-10-218-743-37	Sequence 37, Appli
c 18	722	29.0	1527	9	US-10-218-743-39	Sequence 39, Appli
c 19	722	29.0	1621	9	US-10-218-743-34	Sequence 34, Appli
c 20	722	29.0	1621	9	US-10-218-743-36	Sequence 36, Appli
c 21	638	25.6	449	10	US-09-960-352-678	Sequence 678, App
c 22	476	19.1	1271	9	US-09-923-844B-1	Sequence 1, Appli
c 23	439.5	17.6	1140	9	US-09-938-842A-1684	Sequence 1684, Ap
c 24	439.5	17.6	1140	9	US-09-974-300-646	Sequence 646, App
c 25	405.5	16.3	415	10	US-09-960-352-2589	Sequence 2589, Ap
c 26	402.5	16.1	410	10	US-09-960-352-2589	Sequence 7364, Ap
c 27	402.5	16.1	414	10	US-09-960-352-13050	Sequence 13050, A
c 28	402.5	16.1	424	10	US-09-960-352-3057	Sequence 3057, Ap
c 29	396.5	15.9	418	10	US-09-960-352-3072	Sequence 3072, Ap
c 30	390	15.6	399	10	US-09-960-352-265	Sequence 265, App
c 31	383	15.4	398	10	US-09-960-352-14460	Sequence 14460, A
c 32	381.5	15.3	411	10	US-09-960-352-14919	Sequence 14919, A
c 33	372.5	14.9	384	10	US-09-960-352-5721	Sequence 5721, Ap
c 34	346	13.9	2103	10	US-09-974-300-661	Sequence 661, App
c 35	323	13.0	438	10	US-09-960-352-3315	Sequence 3315, Ap
c 36	319.5	12.8	427	10	US-09-960-352-3315	Sequence 3315, Ap
c 37	319	12.8	422	10	US-09-960-352-5957	Sequence 5957, Ap
c 38	319	12.8	423	10	US-09-960-352-12100	Sequence 12100, A
c 39	316.5	12.7	437	9	US-09-918-995-35463	Sequence 35463, A
c 40	315.5	12.7	427	10	US-09-960-352-3978	Sequence 3978, Ap
c 41	315	12.6	430	10	US-09-960-352-1189	Sequence 1189, Ap
c 42	314	12.6	422	10	US-09-960-352-10642	Sequence 10642, A
c 43	313	12.6	419	10	US-09-960-352-10172	Sequence 10172, A
c 44	312.5	12.5	426	10	US-09-960-352-4979	Sequence 4979, Ap
c 45	311	12.5	419	10	US-09-960-352-4053	Sequence 4053, Ap

ALIGNMENTS

RESULT 1
US-10-004-219B-3
; Sequence 3, Application US/10004219B
; Publication No. US20030087414A1
; GENERAL INFORMATION:
; APPLICANT: Macrozyme
; APPLICANT: Aerts, Johannes M.F.G.
; APPLICANT: Boot, Rolf G.
; TITLE OF INVENTION: A mammalian mucinase, its recombinant production, and
; TITLE OF INVENTION: its use in therapy or prophylaxis against diseases in
; FILE OF INVENTION: which mucus is involved or infection diseases
; FILE REFERENCE: 2183-5136US
; CURRENT APPLICATION NUMBER: US/10/004,219B
; CURRENT FILING DATE: 2001-11-02
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1525
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Combined DNA/RNA Molecule: mouse
; OTHER INFORMATION: AMCase cDNA sequence and deduced amino acid
; OTHER INFORMATION: sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: mouse AMCase
; OTHER INFORMATION: cDNA sequence and deduced amino acid sequence
; FEATURE:

NAME/KEY: CDS
LOCATION: (1) (1419)
US-10-004-219B-3

Alignment Scores:
Pred. No.: 3.6e-281 Length: 1525
Score: 2493.00 Matches: 452
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-004-219B-9 (1-452) x US-10-004-219B-3 (1-1525)

```
QY 1 TyrAsnLeuIleCysThrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPhe 20
DB 64 TACAACCTGTATATGCTATTTCCACCACTGGGCCCGCATGCGGCCAGGTCTGGGGAGCTTC 123
QY 21 LysProAspAspIleAsnProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet 40
DB 124 AAGCCTGATGACATTAAACCCCTGCTGTGTACTCAGCTGATCTATGCTTGTGCGGATG 183
QY 41 GlnAsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrIleAlaPheAsn 60
DB 184 CAGAACAAATGAGATCACCACCATAGAAATGGAATGATGTTACTCTCTATAAAGCTTTCAAT 243
QY 61 AspLeuIleAsnArgAsnSerIleLeuIleThrLeuLeuAlaIleGlyGlyTrpAsnPhe 80
DB 244 GACTTGAAAACAGGAACAGCAACCTGGAACCTCTGCAATTTGGAGGCTGGAACTTT 303
QY 81 GlyThrAlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheIleThr 100
DB 304 GGAACCTGCTCTTCACTACCATGGTTTCCACTTCTCAGAACCCGACAGCTTCATTACC 363
QY 101 SerValIleLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyr 120
DB 364 TCAGTCATCAAAATTTCTGCGTCAGTATGGTTTGTAGGAGCTGGAGCTGGGAATAC 423
QY 121 ProGlySerArgGlySerProGlnAspIleLeuPheThrValLeuValIleGlyGlu 140
DB 424 CCAGGCTCAGTGGGAGCCCTCTCAGGACAAAGCATCTCTTCACTGTCTGGTGAAGGAA 483
QY 141 MetArgGluAlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThr 160
DB 484 ATGCGTGAGCTTTTGGCAGGAGGCTATTGAGAGCAACAGGCCAGAGCTGATGTTACT 543
QY 161 AlaAlaValAlaGlyGlyIleSerAsnIleGlnAlaGlyTyrGluIleProGluLeuSer 180
DB 544 GCTGCTGTAGCTGGTGGGATTTCCAAATCCAGGCTGGCTATGAGATCCCTGAACTTTCT 603
QY 181 LysTyrLeuAspPheIleHisValMetThrTyrAspLeuHisGlySerThrGluGlyTyr 200
DB 604 AAGTACCTGGATTTTCACTCATGTCATGATATGATCATGCTCATGGCTCTGGGAGGCTAC 663
QY 201 ThrGlyGluAsnSerProLeuTyrIleTyrProThrGluThrGlySerAsnAlaTyrLeu 220
DB 664 ACTGGGAGATAGTCTCTTTACAAATACCTTACTGAGACTGTAGCAATGCCTACCTC 723
QY 221 AsnValAspTyrValMetAsnTyrTrpLysAsnAsnGlyAlaProAlaGluLysLeuIle 240
DB 724 AATGTGGATTATGATCATGAACTATTGGAAGAACAAATGGAGCCCGAGCTGAGAAGCTCAT 783
QY 241 ValGlyPheProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIle 260
DB 784 GTTGGATTCCAGAGTATGGACACACTTCACTCCTGAGAACCCCTCTGATATGGAAT 843
QY 261 GlyAlaProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrp 280
DB 844 GGTGCCCTTACCTCTGATGGCCCTGTGGGCCCTATACAGACAGAGCTGGGTCTGG 903
QY 281 AlaTyrTyrGluIleCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSer 300
DB 904 GCCTACTATGAGATTTGCACCTTTCTGAGAAGTGGAGCCACTGAGCTCTGGGATGCTCC 963
```

```
QY 301 GlnGluValProTyrAlaTyrIleAlaAsnGluTrpLeuGlyTyrAspAsnIleLysSer 320
DB 964 CAAAGAGTGGCTATGCTATTAAGCCCAACAGTGGCTTGGCTATGACATATCAGAGC 1023
QY 321 PheSerValLysAlaGlnTrpLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTrpAla 340
DB 1024 TTCACTGTTAAGGCTCAGTGGCTTAAGCAGAACAAATTTTGGAGGTGCCATGATCTGGGCC 1083
QY 341 IleAspLeuAspAspPheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSer 360
DB 1084 ATTGACCTTGATGACTTCACTGGCTCTTCTGTGATCAGGGGAAAATTTCTCTGACTTCT 1143
QY 361 ThrLeuAsnLysAlaLeuGlyIleSerThrGluGlyCysThrAlaProAspValProSer 380
DB 1144 ACTTTGAACAAGCCCTTGGCATATCCACTGAAGTTGCACAGCTCTCCAGCTGCTTCC 1203
QY 381 GluProValThrProGlySerGlySerGlyGlySerGlySerGlySerGlySerSer 400
DB 1204 GAGCCAGTGACTACTCTCCAGGAAGTGGAGTGGGGTGGGAAGCTCCGGAGGAAGCTCT 1263
QY 401 GlyGlySerGlyPheCysAlaAspLysAlaAspGlyLeuTyrProValAlaAspAspArg 420
DB 1264 GGAGGACAGTGGATTCTGTGCCGCAAAAGCAGATGGCTCTACCTGTGGCAGATGACAGA 1323
QY 421 AsnAlaPheTrpGlnCysIleAsnGlyIleThrTyrGlnGlnHisCysGlnAlaGlyLeu 440
DB 1324 AATGCTTTTGGCAGTGTCATCATGGAATCATACACAGCAGCATTTGTCAGAGGGCTT 1383
QY 441 ValPheAspThrSerCysAsnCysCysAsnTrpPro 452
DB 1384 GTTTTGTATACAGCTGTAATTGCTGCAACTGGCCA 1419
```

RESULT 2

US-10-004-219B-2
; Sequence 2, Application US/10004219B
; Publication No. US20030087414A1
; GENERAL INFORMATION:
; APPLICANT: Macrozyme
; APPLICANT: Aerts, Johannes M.F.G.
; APPLICANT: Boot, Rolf G.
; TITLE OF INVENTION: A mammalian mucinase, its recombinant production, and
; TITLE OF INVENTION: its use in therapy or prophylaxis against diseases in
; TITLE OF INVENTION: which mucus is involved or infection diseases
; FILE REFERENCE: 2183-5136US
; CURRENT APPLICATION NUMBER: US/10/004,219B
; CURRENT FILING DATE: 2001-11-02
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2

LENGTH: 1625

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Combined DNA/RNA Molecule: human

OTHER INFORMATION: AMCase cDNA sequence and deduced amino acid

OTHER INFORMATION: sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: human AMCase

OTHER INFORMATION: cDNA sequence and deduced amino acid sequence

FEATURE:

NAME/KEY: CDS

LOCATION: (104) .. (1531)

US-10-004-219B-2

Alignment Scores:

Pred. No.: 5.63e-234 Length: 1625
Score: 2089.50 Matches: 369
Percent Similarity: 90.09% Conservatives: 40
Best Local Similarity: 81.28% Mismatches: 42
Query Match: 83.81% Indels: 3
DB: 9 Gaps: 1

Db 122 GGATTCCTTACCTATGACACAACTTCTCTGAGCAACCCCTCCACACACTCGGAATGGT 181
Qy 263 AlaProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrpAla 281
Db 182 GCGCCACCTCTGGTGTCTGCTGGGCGCTATGCCAAGAGTCTGGGATCTGGGCT 241
Qy 282 TyrTyrGluLeuPheLeuArgSerGlyAlaThrGluValTrpAspAlaSerGln 301
Db 242 TACTACGAGATCTGTACCTTCTGAAAAATGAGCCACTCAGGGATGGATGCCCTCAG 301
Qy 302 GluValProTyrAlaTyrIlysalanGluTrpLeuGlyTyrAspAlaIlysalanPhe 321
Db 302 GAAGTGCCTTATGCTATCAGGCGCAATGTGGTGGCTATGACAAAGCTCAAGAGCTTC 361
Qy 322 SerValLysAlaGlnTrpLeuLysGlnAnanPheGlyGlyAlaMetIleTrpAlaIle 341
Db 362 GATATTAGGCTCAATGGCTTAAGCAACAATCTGGAGGGCGCATGTCTGGGCCATT 421
Qy 342 AspLeuAspAspPheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSerThr 361
Db 422 GATCTGGATGACTTCTACCTGGCACTTCTGCAACAGGCGCAAGTTTCCCTAATCTCCACC 481
Qy 362 LeuAnLysAlaLeuGlyIleSerThrGluGlyCysThrAlaProAspValProSerGlu 381
Db 482 CTGAAGAAGGCGCTCGGCTGAGAGTGAAGTTGCACGCTCCAGCTCAGCCATTGAG 541
Qy 382 ProValThrThrProPro-----GlySerGlySerGlyGlyGlySerGlyGly 398
Db 542 CCAATTAACCTGCTCCAGTGGAGCGGGAACGGAGCGGAGTAGTAGCTCTGGAGGC 601
Qy 399 SerSerGlyGlySerGlyPhe---CysAlaAspLysAlaAspGlyLeuTyrProValAla 417
Db 602 AGCTCGGAGGAGGAGTGGATCTTGTGCTGGCAGACCAACGAGCTTAACCCCTGGGC 661
Qy 418 Asp----AspArgAsnAlaPheTrp-GlnCysIleAsnGlyIleThrTyrGln-GlnHis- 435
Db 662 AAATTAACCAAGAGAGTCCCTCTCGGCACTGCTGTAAGTGAAGTCAAGTACAGGCAAGT 721
Qy 436 CysGlnAlaGlyLeuVal-PheAspThrSerCysAsnCysCysAsnTrp 451
Db 722 TGCCAGGCGGGCTGTCTCTCGACACAGCTGTGAATGCTGCAACTGG 770

RESULT 5

US-09-215-077A-4
; Sequence 4, Application US/09215077A
; Patent No. US20020031793A1
; GENERAL INFORMATION:
; APPLICANT: PRICE, PAUL A.
; APPLICANT: JOHANSEN, JULIA S.
; TITLE OF INVENTION: ASSAY FOR YKL-40 AS A MARKER FOR DEGRADATION OF
; TITLE OF INVENTION: MAMMALIAN CONNECTIVE TISSUE MATRICES
; FILE REFERENCE: 4077-895411US
; CURRENT APPLICATION NUMBER: US/09/215,077A
; PRIOR FILING DATE: 1998-12-18
; PRIOR APPLICATION NUMBER: 08/581,527
; PRIOR FILING DATE: 1996-04-17
; PRIOR APPLICATION NUMBER: 08/089,989
; PRIOR FILING DATE: 1993-07-09
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1681
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:cDNA nucleotide
; OTHER INFORMATION: sequence for the coding region of the gene for
; OTHER INFORMATION: YKL-40.
US-09-215-077A-4

Alignment Scores:
Pred. No.: 1.01e-102 Length: 1681
Score: 966.50 Matches: 187

Percent Similarity: 66.07% Conservative: 70
Best Local Similarity: 48.07% Mismatches: 123
Query Match: 38.77% Indels: 10
DB: 10 Gaps: 5
US-10-004-219B-9 (1-452) x US-09-215-077A-4 (1-1681)
Qy 1 TyrAsnLeuIleCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPhe 20
Db 135 TACAACTGGTCTGCTACTACACAGCTGGTCCCAGTACCGGAGGCGGATGGGAGCTGC 194
Qy 21 LysProAspAlaIleAsnProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet 40
Db 195 TTCCAGATGCCTTGACCGCTTCTGTGTATCCACATCATCATCAGCTTTGCCAATATA 254
Qy 41 GlnAnanGluIleThrThrIleGluTrpAnanPheValThrLeuTyrLysAlaPheAn 60
Db 255 AGCAACGATCATCGACACCTGGAGTGGAAATGATGTAGCGCTCTACGGCATGCTCAAC 314
Qy 61 AspLeuLysAsnArgAsnSerLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAnanPhe 80
Db 315 ACATCTCAACACACGACCCCACTGAAGACTCTTGTCTGCGAGGATGGAACTTT 374
Qy 81 GlyThrAlaProPheThrThrMetValSerThrSerGlnAnanArgGlnThrPheIleThr 100
Db 375 GGTCTCAAGATTTTCCAGATAGCTCCAAACACCCAGAGTCCGCGACTTTTCATCAAG 434
Qy 101 SerValIleLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyr 120
Db 435 TCAGTACCGCCATTCTCGCCACCCCTGCTTGTATGGCGGAGCTTGCCTGGGCTCTAC 494
Qy 121 ProGlySerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValLysGlu 140
Db 495 CTGAGCGGAGA-----GACAACACCATTTTACCACCTAATCAAGGNA 539
Qy 141 MetArgGluAlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThr 160
Db 540 ATGAAGCGCGAATTTATAAAGAGAGCC--CAGCCAGCGGAAAAAGCAGCTCTCGTCCAG 596
Qy 161 AlaAlaValAlaGlyIlyleSerAnanIleGlnAlaGlyTyrGluIleProGluLeuSer 180
Db 597 GCAGCACTGTCTGCGGGGGAAGGTCAACCATTCAGCAGCATGATGACATTCGCCAAGATATCC 656
Qy 181 LysTyrLeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr 200
Db 657 CAACACCTGGATTTTATTAGCATCATGACCTACGATTTTTCATGGCGCTGGCGTGGGACC 716
Qy 201 ThrGlyGluAnanSerProLeuTyrIlysalanTyrProThrGluThrGlySerAsnAlaTyr 220
Db 717 ACAGGCCATCAGTCCCTCAGGCGAGGTTCAGGAGGATGCAAGTCTGACAGATTTCAGC 776
Qy 221 AnValAspTyrValMetAsnTyrTrpLysAnanGlyAlaProAlaGluLysLeuIle 240
Db 777 AACACTGACTATGCTGTGGGGTACATGTTGAGGCTGGGGGCTCTCGCCAGTAAAGTGGTG 836
Qy 241 ValGlyPheProGluTyrGlyHisThrPheIleLeuArgAnanProSerAspAnGlyIle 260
Db 837 ATGGGCATCCCACTTCGGAGGAGGACTTCACTCTGGCTTCT--TCTGAGACTGGTGT 893
Qy 261 GlyAlaProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrp 280
Db 894 CCAGCGCCAACTCAGGACCGGGAATTCAGGCGGTTCCAGCGGTTCCACCAAGGAGCGAGCCCTT 953
Qy 281 AlaTyrTyrGluIleCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSer 300
Db 954 GCTACTATGAGATCTGTGACTTCTCTCGC---GGAGCCACAGTCCATAGAACCTCGGC 1010
Qy 301 GlnGluValProTyrAlaTyrIlysalanGluTrpLeuGlyTyrAspAnanIlysalan 320
Db 1011 CAGCAGGTCCTCTATGCCACCAAGGCGCAACAGTGGTAGGATACGACGACGAGAAAGC 1070
Qy 321 PheSerValLysAlaGlnTrpLeuLysGlnAnanPheGlyGlyAlaMetIleTrpAla 340

Db 1071 GTCAAGCAAGTGGTACCTGTAAGGATAGGAGTGGCAGCCCATGGTATGGGCC 1130
QY 341 ILeAspLeuAspPheThrGlySerPheCysAspGlnGly---LysPheProLeuThr 359
Db 1131 CTGACCTGGATGACTTCCAGGCTCTTCTGCGCAGGATCTGCGTCTCTCACC 1190
QY 360 SerThrLeuAsnLysAlaLeuGlyIleSerThrGluGlyCysThrAlaProAspValPro 379
Db 1191 AATGCCATCAAGGATGACTCGCTGC-AAGGTAGCCCTCTGTTCTGCACACAGCAGCGGG 1249
QY 380 SerGluProValThrThrProProGly 388
Db 1250 GCCAAGGATGCCCGCTCCCGCTGTC 1276

RESULT 6
US-09-262-213A-4
; Sequence 4, Application US/09262213A
; Patent No. US20020090658A1
; GENERAL INFORMATION:
; APPLICANT: PRICE, PAUL
; APPLICANT: JOHANSEN, JULIA
; TITLE OF INVENTION: ASSAY FOR YKL-40 AS A MARKER FOR DEGRADATION OF
; TITLE OF INVENTION: MAMMALIAN CONNECTIVE TISSUE MATRICES
; FILE REFERENCE: 407T-895412US
; CURRENT APPLICATION NUMBER: US/09/262,213A
; PRIOR FILING DATE: 1999-03-04
; PRIOR APPLICATION NUMBER: US 08/089,989
; PRIOR FILING DATE: 1993-07-09
; PRIOR APPLICATION NUMBER: PCT/US94/07754
; PRIOR FILING DATE: 1994-07-08
; PRIOR APPLICATION NUMBER: US 08/581,527
; PRIOR FILING DATE: 1996-04-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 1681
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-262-213A-4

Alignment Scores:
Pred. No.: 1,01e-102 Length: 1681
Score: 966.50 Matches: 187
Percent Similarity: 66.07% Conservative: 70
Best Local Similarity: 48.07% Mismatches: 123
Query Match: 38.77% Indels: 10
DB: 10 Gaps: 5

US-10-004-219B-9 (1-452) x US-09-262-213A-4 (1-1681)

QY 1 TyrAsnLeuLeuCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPhe 20
Db 135 TACAACTGGTCTGCTACTACACAGCTGGTCCAGTACCGGAGGAGGATGGAGCTGC 194
QY 21 LysProAspAspIleThrThrLeuGlyCysThrHisLeuIleTyrAlaPheAlaGlyMet 40
Db 195 TTCCACAGATGCCCTTGACCGCTCTCTGTGTGTACCCACATCATCTACAGCTTTCGAATAFA 254
QY 41 GlnAsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrLysAlaPheAsn 60
Db 255 AGCAACGATCATCATCACACACCTGGGAGTGGAGTATGATGACGCTCTACGGCATGCTCAAC 314
QY 61 AspLeuLysAsnArgAsnSerLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPhe 80
Db 315 ACATCAACACACAGNACCCACCTGAGATCTCTTGTCTGTGGAGATGGAACTTT 374
QY 81 GlyThrAlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheIleThr 100
Db 375 GGCTCTCAAGAGATTTTCCAAGATAGCTCCACACACAGGAGTGGCGGACTTTTCATCAAG 434
QY 101 SerValIleLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyr 120
Db 435 TCAGTACCCGCCATTTCTGCGCACCCCATGGCTTGTATGGGCGTGACCTTCCCTGCTTAC 494

QY 121 ProGlySerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValLysGlu 140
Db 495 CTGGACGGAGA-----GACAAACACCATTTTACCACCTAATCAAGGAA 539
QY 141 MetArgGluAlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThr 160
Db 540 ATGAAGGCGCAATTTATAAAGGAGCC---CAGCCAGGAGGAAAAAGCAGCTCTGCTCAGC 596
QY 161 AlaAlaValAlaGlyGlyLysSerAsnIleGlnAlaGlyTyrGluIleProGluLeuSer 180
Db 597 GCAGCACTGTCTGCGGGGAAGGTCAACATTGACAGCAGCTATGACATTGCCAAGATATCC 656
QY 181 LysTyrLeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTrpGluLysTyr 200
Db 657 CAACACCTGGATTTTATTAGCATCATGACCTACGATTTTTCATGGCGCTGGCGTGGACC 716
QY 201 ThrGlyGluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeu 220
Db 717 ACAGGCCATCACAGTCCCTCAGCGCAGGTCAAGGAGTCAAGTCTGCACAGATTTCAGC 776
QY 221 AsnValAspTyrValMetAsnTyrTrpLysAsnAsnGlyAlaProAlaGluLysLeuIle 240
Db 777 AACACTGACTATGCTGTGGGTACATGTTGAGGCTGGGGGCTCTGCCAAGTAAGCTGGT 836
QY 241 ValGlyPheProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIle 260
Db 837 ATGGGCATCCCACTTCGGGAGGAGCTTCACTCTGGCTTCT---TCTGAGACTGGTGT 893
QY 261 GlyAlaProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrp 280
Db 894 CCAGGCCAATCTCAGGACCGGGAATTCAGGCCGCTTCCAGAGGAGGAGGAGGAGCCCTT 953
QY 281 AlaTyrTyrGluIleCysThrPheLeuArgSerGlyAlaThrGluValTrpAspAlaSer 300
Db 954 GCCTACTATGAGATCTGTGACTTCTCTCCGC---GGAGCCACAGTCCATAGAACCTCCGC 1010
QY 301 GlnGluValProTyrAlaTyrLysAlaAsnGlnTrpLeuGlyTyrAspAsnIleLysSer 320
Db 1011 CAGCAGTCCCTTATGCCACCAAGGCAACCACTGGTGGTAGGATAGCAGCAGCAGGAAAGC 1070
QY 321 PheSerValLysAlaGlnTrpLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTrpAla 340
Db 1071 GTCAAAAGCAAGGTGCAGTACCTGAAAGATAGGAGCAGCTGGCAGGCGCCATGATGGGCC 1130
QY 341 IleAspLeuAspAspPheThrGlySerPheCysAspGlnGly---LysPheProLeuThr 359
Db 1131 CTGGACCTGGATGACTTCCAGGCTCTCTTCTGCGGCCAGGATCTGCGCTTCCCTCTCACC 1190
QY 360 SerThrLeuAsnLysAlaLeuGlyLysSerThrGluGlyCysThrAlaProAspValPro 379
Db 1191 AATGCCATCAAGGATGACTCGCTGC-AACGTAGCCCTCTGTTCTGTCACACAGCAGCGGG 1249
QY 380 SerGluProValThrThrProProGly 388
Db 1250 GCCAAGGATGCCCGCTCCCGCTGTC 1276

RESULT 7

US-09-822-830A-402
; Sequence 402, Application US/09822830A
; Patent No. US20020142952A1
; GENERAL INFORMATION:
; APPLICANT: Genetics Institute, Inc.
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fechtel, Kim
; APPLICANT: Agostino, Michael J.
; APPLICANT: Howes, Steven H.
; APPLICANT: Resnick, Richard J.
; APPLICANT: Gulukota, Kamalakar
; APPLICANT: Graham, James R.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
; FILE REFERENCE: GIN 6402

DB: 10 Gaps: 6
US-10-004-219B-9 (1-452) x US-09-765-231A-8 (1-1474)
QY 1 TyrAsnLeuIleCysTyrPheThrAsnTrp-AlaGlnTyrArgProGlyLeuGlySerPh 20
DB 137 TACAAACTGGTTGGCTTACCACTGGTCCAGGACCGGACGGAACCAAGGAAAT 196
QY 20 eLysProAspAsp-IleAsnProCysLeuCysThrHisLeuIleTyrAlaPhe-AlaGly 39
DB 197 CACCCCTGAGGATATGACCCCTTCTATGCTCTCATCTCATCTATTCATTCGCGCAGC 256
QY 40 MetGlnAsnGlnIleThrThrIleGluTrpAsnAspValThrLeuTyrLysAlaPhe 59
DB 257 ATCGAAACCAACAGGTTATATCAAGGACAGAGTGAAGTGTCTCTACACAGACCATC 316
QY 60 AsnAsp-LeuLysAsnArgAsnSerLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAs 79
DB 317 AACAGTTCTCAAAACCAAGAAATCCCAAACTGAAATTTCTTTGTCCATTTGGAGGTACT 376
QY 79 nPheGlyThrAlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheI 99
DB 377 GTTGTGTTCAAAAGGTTTCCACCTATGTTGATGTTCTTCTACATCAGCTTGAATTCAT 436
QY 99 eThrSerValIleLeuPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpG 119
DB 437 TAACTCCATTAATCTGTTCTGAGGAACATAACTTTGATGGACTGGATGTAAGCTGGAT 496
QY 119 uTyrProGlySerArgGlySerProProGlnAspLysHisLeuPheThrValLeuVally 139
DB 497 CTACCCAGATCAGAA-----GAAACACACTCAT---TTCACTGTCTGATTCA 541
QY 139 sGluMetArgGluAlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMetVa 159
DB 542 TGAATTAGCAGAGGCTTTTCAGAGGACTTCACAAATCCACCAAGGAAGGCTTCTCT 601
QY 159 lThrAla-AlaValAlaGlyIleSerAsnIleGlnAlaGlyTyrGluIleProGluL 179
DB 602 GACTGGGGGGGTATCTGCAGGAGGCAAAATGATTGATAACAGCTATCAAGTTGAGAAC 661
QY 179 eSerLysTyrLeuAspPheIleHieValMetThrTyrAspLeuHisGlySerTrpGluG 199
DB 662 TGCAGAAAGATCTGGATTTTCATCACTCTCTCTTTCATCTTCACTTCCATGGTCTTGGAA 721
QY 199 ly-----TyrThrGlyGluAsnSerPro-LeuTyrLysTyrProThrGluThrGlySer 216
DB 722 AGCCCTTATCTAGTCCCAACAGCCCTCTGAGCAAGGGTGGCAGGACGAGGGCCA 781
QY 217 AsnAlaTyrLeuAsnValAspTyrValMetAsnTyrTrpLysAsnGlyAlaProAla 236
DB 782 AGTCTCTACTACATGTGGAATATGCTGGGGTACTGGATACATAAGGGAATGCCATCA 841
QY 237 GluLysLeuIleValGlyPheProGluTyr-GlyHisThrPheIleLeuArgAsnProSe 256
DB 842 GAGAGGTGGTCATGGGCAATCCCATATGGGCACTCTTTCACACTGGCCTCT---GC 898
QY 256 rAspAsnGlyIleGlyAlaProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgL 276
DB 899 AGAACCCCGTGGGGGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 958
QY 276 nAlaGlyPheTrpAlaTyrTyrGluIleCysThrPheLeuArgSerGlyAlaThrGluVa 296
DB 959 TTCAGGCTTCTGGCCCTATTATGAGATCTGCCAGTTCTCTGAAA---GGAGCAAGATCAC 1015
QY 296 lTrpAspAlaSerGlnGluValProTyrAlaTyrLysAlaAsnGluTrpLeuGlyTyrAs 316
DB 1016 GCGGCTCCAGGATCAGCAGGTTCCCTTACGAGTCAAGGGGAAACCAAGTGGGTGGCTATGA 1075
QY 316 pAsnIleLysSerPheSerValLysAlaGlnTrpLeuLysGlnAsnAsnPheGlyGlyAl 336
DB 1076 TGATGGAAGATGAGGACCAAGGTTTCAGTTCTTAAAGAAATTTAAACCTCGGAGGAC 1135
QY 336 aMetIleTrpAlaIleAspLeuAspAspPheThrGlySerPheCysAspGlnGlyLysPh 356

DB 1136 CATGATCTGGTCTATTGACATGGATGACTTCACTGGCAAAATCTCTCAACACGAGGCCCTTA 1195
QY 356 eProLeuThrSerThrLeuAsnLysAlaLeuGly-----IleSerThrGluG 372
DB 1196 CCCTTGTTCACAGCAGTCAAGAGAGGCTTGGCTCCCTGTGAAGGATTAACCTACAGAG 1255
QY 372 yCysThrAlaProAspValProSerGluProValThrThrProProGlySerGlySerG 392
DB 1256 AAGCAGGCAAGATGACCTTGTCTGCTGGGGCTGCTCTCTCCAGGAATTTCTCATGTGG 1315
QY 392 Y 392.
DB 1316 A 1316

RESULT 9
US-10-218-743-20
; Sequence 20, Application US/10218743
; Publication No. US20030096779A1
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
; FILE REFERENCE: AL-2-C3
; CURRENT APPLICATION NUMBER: US/10/218,743
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: US/09/292,225
; PRIOR FILING DATE: 1999-04-15
; PRIOR APPLICATION NUMBER: 60/098,909
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/085,295
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/098,565
; PRIOR FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: 09/062,013
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 1608
; TYPE: DNA
; ORGANISM: Dermatophagoides farinae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1608)
US-10-218-743-20

Alignment Scores:
Pred. No.: 4,698-75 Length: 1608
Score: 729.50 Matches: 154
Percent Similarity: 53.61% Conservative: 69
Best local Similarity: 37.02% Mismatches: 144
Query Match: 29.26% Indels: 49
DB: 9 Gaps: 11

US-10-004-219B-9 (1-452) x US-10-218-743-20 (1-1608)
QY 3 LeuIleCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPheLysPro 22
DB 46 ATTGTTTCTTATGTTGAACATCGTCCGTATATCATATAA---GTTGATCCATACACTATC 102
QY 23 AspAspIleAsnProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyMetGlnAsn 42
DB 103 GAAGATTTGATTCATTCACAGGTACACATTAAATGATGTTTTCGCTAAATTCATGAA 162
QY 43 AsnGluIleThrThr-----IleGluTrpAsn 51
DB 163 TACAAATACACAAATCAAGTTTTCGATCTTCCACAGATGATACCAATACATG--- 219
QY 52 AspValThrLeuTyrLysAlaPheAsnAspLeuLysAsnArgAsnSerLysLeuLysThr 71

220 GAAAAAGCTGGTTATGAACCTGTTTCAACCAACTTGGCATTGAAGATCAGAAATTAAACCAACC 279
72 LeuLeuAlaIleGlyGlyTyrPAsnPheGlyThrAlaProPheThrThrMetValSerThr 91
ATGATTTCACTTGGTGGTGGTGAAGGCTCGGAAAAATATTCCGATATGGCTGCAAAAT 339
92 SerGlnAsnArgGlnThrPheIleThrSerValIleLysPheLeuArgGlnTyrGlyPhe 111
340 CCAACATATCGTCAACAATTCATACATCAGTTTGGACTTTTTCGAAGAATACAAAGTTTC 399
112 AspGlyLeuAspLeuAspTyrGluTyrProGlySerArg--GlySerProProGlnAsp 130
400 GACGGCTAGATTTGGATTTGGAGATCTCTGGATCTCGATTTGGTGTACCCGAAAAATCGAT 459
131 LysHisLeuPheThrValLeuValLysGluMetArgGluAlaPheGluGlnGluAlaIle 150
460 AAACAAACTATTGGCTTTGGTTAGAGAACTTAAAGACGCTTTTGAACCTCATGGC-- 516
151 GluSerAsnArgProArgLeuMetValThrAlaAlaValAlaGlyGlyIleSerAsnIle 170
517 -----TACTTTGTTCAGTGTGCAGTATCACACAGTAAAGACAAATC 558
171 GlnAlaGlyTyrGluIleProGluLeuSerLysTyrLeuAspPheIleHisValMetThr 190
559 GACCGACTTATGATATCAAGAATTTGAACAAAATTTGCGATTTGGATGAATGTCATGACA 618
191 TyrAspLeuHisGlySerTyrGluGlyTyrThrGlyGluAsnSerProLeuTyrLysTyr 210
619 TATGATTACACCGTGGATGGGAAAACCTTTACGGTCACAATGCTCGTTGTATAAACA 678
211 ProThrGluThrGlySer---AsnAlaTyrLeuAsnValAspTyrValMetAsnTyrTyr 229
679 CCAGATGAAACTGATGATGTTGCACACTTACTTCAATGTCAACTACACCATGCACTATTAT 738
230 LysAsnAsnGlyAlaProAlaGluLysLeuIleValGlyPheProGluTyrGlyHisThr 249
739 TTGAACAATGGTGCACACAGACACAAATTTGGTATGGTGTTCATTTCTATGGCCGGTGT 798
250 PheIleLeuArgAsnProSerAspAsnGlyIleGlyAlaProThrSerGlyAspGlyPro 269
799 TGGAGCATTTGAAGATCGAAGCAAACTCAAACTTGGAGATCCAGCAAGGCATGTCGCC 858
270 AlaGlyAlaTyrThrArgGlnAlaGlyPheThrAlaTyrTyrGluIleCysThrPheLeu 289
859 CCAGTTTCATTTCTGGTGAAGAAGGTGCTCTCATATAGAAATTTGTGTCAAATTTGTT 918
290 ArgSerGlyAlaThrGluValTyr-----AspAlaSerGlnGluValProTyr 305
919 CAAAA-----GAAGAATGGCATATCCAAATPACGATGAATATTACAATGCTCCATAT 969
306 AlaTyrLysAlaAsnGluTyrLeuGlyTyrAspAsnIleLysSerPheSerValLysAla 325
970 GGTTCAAATGATAAATCTGGTTCGTTACATGATCTGGCCAGTATATCATGCAAGTTG 1029
326 GlnTyrLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTyrAlaIleAspLeuAspAsp 345
1030 GCTTTCCTCGAAAGAATTAGGCGTTCTTGGTGTCTGATGTTGGTCATTTGGTGAATAATGAT 1089
346 PheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSerThrLeuAsnLysAla 365
1090 TTCAAAGT---CACTCGGA-----CCGAAAAATCCATTTGTTGAACAAAAGTT 1134
366 -----LeuGlyIleSerThr 370
1135 CATAAATATGATTAATGGCGATGAAAGAAGCTCTTTCGAATGCAATTTGGGTGCCAAGTACA 1194
371 GluGlyCysThrAlaProAspValProSerGluProValThrThrPro 386
1195 ACGACACCAACTCCACACGACACACCCACACCCCGACTACACGCCA 1242

```
; Publication No. US2003009679A1
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: AU-2-C3
; CURRENT APPLICATION NUMBER: US/10/218,743
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: US/09/292,225
; PRIOR FILING DATE: 1999-04-15
; PRIOR APPLICATION NUMBER: 60/098,909
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/085,295
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/098,565
; PRIOR FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: 09/062,013
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 1608
; TYPE: DNA
; ORGANISM: Dermatophagoides farinae
US-10-218-743-22

Alignment Scores:
Pred. No.:      4.69e-75      Length:      1608
Score:          729.50       Matches:     154
Percent Similarity: 53.61%   Conservative: 69
Best Local Similarity: 37.02% Mismatches:    144
Query Match:     29.26%     Indels:       49
DB:              9          Gaps:        11

US-10-004-219B-9 (1-452) x US-10-218-743-22 (1-1608)

QY      3  LeuileCvstYrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPhelysPro 22
Db      1563 ATGTGTTGGTATTGTTGGACATGTCGGTATATCATAAA---GTTGATCCATACACTATC 150
QY      23  AspAspileAsnProCysLeuCyeThrHisleuleTyrAlaPheAlaGlyMetGlnAsn 42
Db      1506 GAAGATATTGATCCATTCAAGTGTCACACATTAAATGATGTTGCTTGAATAATTGCATGAA 144
QY      43  AsnGluileThrThr-----IleGluTrpAsn 51
Db      1446 TACAATAATCAATTCAGTTTCGATCTTACCAGATGATAACCAATCACTCATGG--- 139
QY      52  AspValThrLeuTyrLysAlaPheAsnAspLeuLysAsnArgAsnSerLysLeuLysThr 71
Db      1389 GAATAACGTGGTATGAACGTTTCACAACTGGATTGGAATCCAGAATCCAGATTAACCACC 133
QY      72  LeuLeuAlalaileGlyGlyTyrAsnPheGlyThrAlaPropheThrThrMetValSerThr 91
Db      1329 ATGATTTCACTTGGTGGTGGTATGATGAGGCTCGGAAAAAATATTCGATATGGCTGCAAA 127
QY      92  SerGlnAsnArgGlnThrPheIleThrSerVallelleLysPheLeuArgGlnTyrGlyPhe 111
Db      1269 CCAACATATCGTCAACAAATTCATACAATCAGTTTTGGGACTTTTTGCAAGAAATACAA 121
QY      112  AspGlyLeuAspLeuAspTrpGluTyrProGlySerArg---GlySerProProGlnAsp 130
Db      1209 GACGGCTTAGATTGGATTGGAGATTCCTGGATCTCGATTGGGGTAACCCGAAAAATCGAT 115
QY      131  LysHisLeuPheThrValLeuVallyLysGluMetArgGluAlaPheGlnGlnAlaIle 150
Db      1149 AAACAAACTATTTTGGCTTTGGTTAGAGAACTTAAGACGCTTTTGAACCTCATGGC--- 109
QY      151  GluSerAsnArgProArgLeuMetValThrAlaAlaValAlaGlyGlyIleSerAsnIle 170
Db      1092 -----TACTGTTCACCTGCTGAGTATCACCAAGGTAAAGACAAAATC 105
```

```
171 GlnAlaGlyTyrGluLeuProGluLeuSerLysTyrLeuAspPheIleHisValMetThr 190
1050 GACCGAGCTTATGATCAAGAAATGAACAATTTGTCGATGGATGATCATGACA 991
191 TyrAspLeuHisGlySerTrpGluGlyTyrThrGlyGluAsnSerProLeuTyrLysTyr 210
990 TATGATTACCAGTGGATGGGAAACTTTTACGGTGCACAAATGCTCCGTTGTATAAACA 931
211 ProThrGluThrGlySer---AsnAlaTyrLeuAsnValAspTyrValMetAsnTyrTrp 229
930 CCAGATGAACATGAGTGGACATCTTCAATGTCAACTACACCAATGACATATAT 871
230 LysAsnAsnGlyAlaProAlaGluLysLeuLeuValGlyPheProGluTyrGlyHisThr 249
870 TTGAACATGGTGGCCACAGACAAATTTGAATGGTGTCCATCTATGGCCGTGT 811
250 PheIleLeuArgAsnProSerAspAsnGlyIleGlyAlaProThrSerGlyAspGlyPro 269
810 TGGAGCATTTGAAGATCGAAGCAAACTCAAACTGGAGATCCAGCCAAAGGCATGCGCC 751
270 AlaGlyAlaTyrThrArgGlnAlaGlyPheTrpAlaTyrGluIleCysThrPheLeu 289
750 CCAGGTTTCATTTCTGGTGAAGAAGTGTCTCTCATATATAGAAATGTGTCAATGT 691
290 ArgSerGlyAlaThrGluValTrp-----AspAlaSerGlnGluValProTyr 305
690 CAAAAA-----GAAGATGGCATATCCAAATACGATGAATATACATGCTCCATAT 640
306 AlaTyrLysAlaAsnGluTrpLeuGlyTyrAspAsnIleLysSerPheSerValLysAla 325
639 GGTTCATATGATAAAATCTGGTGGTTCAGATGATCGGCCAGTATATATACGAAATG 580
326 GlnTrpLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTrpAlaIleAspLeuAsp 345
579 GCTTTCTCTCAAAAGAAATAGCGTTCTGTGTGTCATGTTGGTTCATGGAAATGATGAT 520
346 PheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSerThrLeuAsnLysAla 365
519 TTCAAAGGT---CACTGCGGA-----CCGAAAAATCCATTTGTGAACAAAGTT 475
366 -----LeuGlyIleSerThr 370
474 CATATATGATTAAATGGCGATGAAAGAACTCTTTGCAATGCAATTTGGTTCAGATACA 415
371 GluGlyCysThrAlaProAspValProSerGluProValThrPro 386
414 ACGACACCAACTCCAAACGACGACACCCCAACCCCGACTACAAACGCCA 367
```

RESULT 11

```
US-10-218-743-17
; Sequence 17, Application US/10218743
; Publication No. US20030096779A1
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
; AND USES THEREOF
; FILE REFERENCE: AL-2-C3
; CURRENT APPLICATION NUMBER: US/10/218,743
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: US/09/292,225
; PRIOR FILING DATE: 1999-04-15
; PRIOR APPLICATION NUMBER: 60/098,909
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/085,295
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/098,565
; PRIOR FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: 09/062,013
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 49
```

```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 1665
; TYPE: DNA
; ORGANISM: Dermatophagoides farinae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1665)
US-10-218-743-17

Alignment Scores:
Pred. No.: 4,95e-75 Length: 1665
Score: 729.50 Matches: 154
Percent Similarity: 53.61% Conservative: 69
Best Local Similarity: 37.02% Mismatches: 144
Query Match: 29.26% Indels: 49
DB: 9 Gaps: 11
```

US-10-004-219B-9 (1-452) x US-10-218-743-17 (1-1665)

```
QY 3 LeuIleCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPheLysPro 22
DB 103 ATGTTTGTATGTTGGAAACATGTCCTGATATATCAAAA---GTTGATCCATACACTATC 159
QY 23 AspAspIleAsnProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyMetGlnAsn 42
DB 160 GAAGATATTGATCCATTCAAGTGTACACATTAAATGATGTTTCGCTTAAATTTGATGAA 219
QY 43 AsnGluIleThrThr-----IleGluTrpAsn 51
DB 220 TACAAATACACAATTCAAAGTTTTCGATCTTACCAGATGATTAACATAACTCATGG--- 276
QY 52 AspValThrLeuTyrLysAlaPheAsnAspLeuLysAsnArgAsnSerLysLeuLysThr 71
DB 277 GAAACACGTGGTTATGAACGTTTCAACAACCTGCGATGGAAGATCCAGAAATTAACACC 336
QY 72 LeuLeuAlaIleGlyGlyTyrAsnPheGlyThrAlaProPheThrMetValSerThr 91
DB 337 ATGATTTCACCTGGTGGTTGGTATGAAGCTCGGAAAAATATATCCGATATGCTGCAAA 396
QY 92 SerGlnAsnArgGlnThrPheIleThrSerValIleLysPheLeuArgGlnTyrGlyPhe 111
DB 397 CCAACATATCGTCAACAATTCATAACATCAGTTTGGACTTTTTCGAAGAATACAAAGTTC 456
QY 112 AspGlyLeuAspLeuAspTrpGluTyrProGlySerArg---GlySerProGlnAsp 130
DB 457 GACGGTCTAGATTGGATGGAGATTCCTGGATCTCGATTGGGTAAACCGGAAATTCGAT 516
QY 131 LysHisLeuPheThrValLeuValLysGluMetArgGluAlaPheGluGlnGluAlaIle 150
DB 517 AAACAAAACTATTTGGCTTTGGTAGAGAACTTAAAGACGCTTTTGAACCTCATGGC--- 573
QY 151 GluSerAsnArgProArgLeuMetValThrAlaValAlaAlaGlyGlyIleSerAsnIle 170
DB 574 -----TACTTGTGTGACTGCTGAGTATCACCAGGTAAACAGAAAAATC 615
QY 171 GlnAlaGlyTyrGluIleProGluLeuSerLysTyrLeuAspPheIleHisValMetThr 190
DB 616 GACCGAGCTTATGATATCAAGAATTTGAACAATTTGTCGATGGATGAATGTCATGACA 675
QY 191 TyrAspLeuHisGlySerTrpGluGlyTyrThrGlyGluAsnSerProLeuTyrLysTyr 210
DB 676 TATGATTACCACCGTGGATGGGAAACTTTTACGGTGCACAAATGCTCCGTTGTATAAACA 735
QY 211 ProThrGluThrGlySer---AsnAlaTyrLeuAsnValAspTyrValMetAsnTyrTrp 229
DB 736 CCAGATGAACATGATGATGTCACACTTACTTCAATGTCAACTACACCATGACATATTAT 795
QY 230 LysAsnAsnGlyAlaProAlaGluLysLeuIleValGlyPheProGluTyrGlyHisThr 249
DB 796 TTGAACAATGGTGGCCACAGACAAATTTGAATGGTGTTCATTTCTATGGCCGTGTCT 855
QY 250 PheIleLeuArgAsnProSerAspAsnGlyIleGlyAlaProThrSerGlyAspGlyPro 269
```

Db 856 TGGAGCATTTGAAGATCGAAGCAAACTCAAACCTTGGAGATCCAGCCAAAGCATGTCGCC 915
Qy 270 AlaGlyAlaTyThrArgGlnAlaGlyPheTTPAlaTyThrGluLeuLeuLeuLeu 289
Db 916 CCAGGTTTCATTTCTGGTGGAGAGTGTCTCTCATATATAGAAATGTGCAATGTTT 975
Qy 290 ArgSerGlyAlaThrGluValTTP-----AspAlaSerGlnGluValProTyr 305
Db 976 CAAAAA-----GAGAATGGCATATCCAAATACGATGAATATATACAAATGCTCCATAT 1026
Qy 306 AlaTyThrLysAlaAsnGluTTPLeuGlyTyThrAspAsnLeuLysSerPheSerValLysAla 325
Db 1027 GGTTCATATGATATAAATCTGGTGGTGTCTGATGATCTGCCAGTATATCATGCAAGTTG 1086
Qy 326 GlnTTPLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTTPAlaLeuAspLeuAsp 345
Db 1087 GCTTTCCTGAAAGAAATAGGCGTTTCTGTGTCATGTTGTGTCATGGAATATGATAT 1146
Qy 346 PheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSerThrLeuAsnLysAla 365
Db 1147 TTCAAAGGT---CACTGCGGA-----CCGAAAAATCCATTGTTGAACAAAGTT 1191
Qy 366 -----LeuGlyLysSerThr 370
Db 1192 CATAATATGATTAATGGCGATGAAAGAACTCTTTTGAATGCAATTTTGGGTCCAAAGTACA 1251
Qy 371 GluGlyCysThrAlaProAspValProSerGluProValThrPro 386
Db 1252 ACGACACCACTCCCAACGACGACCAACCAACCCGCACTACAAACGCCA 1299

RESULT 12

US-10-218-743-19/c
; Sequence 19, Application US/10218743
; Publication No. US20030096779A1
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: AL-2-C3
; CURRENT APPLICATION NUMBER: US/10/218,743
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: US/09/292,225
; PRIOR FILING DATE: 1999-04-15
; PRIOR APPLICATION NUMBER: 60/098,909
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/085,295
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/098,565
; PRIOR FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: 09/062,013
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 1665
; TYPE: DNA
; ORGANISM: Dermatophagoides farinae
US-10-218-743-19

Alignment Scores:

Pred. No.: 4,95e-75 Length: 1665
Score: 729.50 Matches: 154
Percent Similarity: 53.61% Conservative: 69
Best Local Similarity: 37.02% Mismatches: 144
Query Match: 29.26% Indels: 49
DB: 9 Gaps: 11

US-10-004-219b-9 (1-452) x US-10-218-743-19 (1-1665)

Qy 3 LeuIleCysTyThrPheThrAsnTTPAlaGlnTyThrArgProGlyLeuGlySerPheLysPro 22

Db 1563 ATTGTTTGTATTGTTGAACATGGTCCGTATATATATAAA---GTTGATCCATACACTATC 1507
Qy 23 AspAspIleAsnProCysLeuCysThrHisLeuIleTyThrAlaPheAlaGlyMetGlnAsn 42
Db 1506 GAAGATATTGATCCATCAAGTGTACACATTTAATGTATGTTTTCGCTAAAATGATGAA 1447
Qy 43 AsnGluIleThrThr-----IleGluTTPAsn 51
Db 1446 TACAAATACACAATTCAAAGTTTTCGATCCTTACCAAGATGATAACCAATCATCTG--- 1390
Qy 52 AspValThrLeuTyThrLysAlaPheAsnAspLeuLysAsnArgAsnSerLysLeuLysThr 71
Db 1389 GAAAAACGCTGGTATGAACGTTTCAACAACTTGGCATTCGAAGAATCCAGAATTAACCAACC 1330
Qy 72 LeuLeuAlaIleGlyGlyTTPAsnPheGlyThrAlaProPheThrThrMetValSerThr 91
Db 1339 ATGATTTCACTTGTGGTGTGTTGATGAAGCTCGGAAAAAATATTCGATATGCTGCAAT 1270
Qy 92 SerGlnAsnArgGlnThrPheIleThrSerValIleLysPheLeuArgGlnTyThrGlyPhe 111
Db 1269 CCAACATATGTCACAATTCATACAATCAGTTTGGACTTTTTCGAAGAATACAAAGTTT 1210
Qy 112 AspGlyLeuAspLeuAspTTPGluTTPProGlySerArg---GlySerProProGlnAsp 130
Db 1209 GACGGCTAGATTTGGATTGGAGTATCTCGATCTCGATTGGGTAAACCGAAAAATCGAT 1150
Qy 131 LysHisLeuPheThrValLeuValLysGluMetArgGluAlaPheGluGlnGluAlaIle 150
Db 1149 AAACAAACTATTTGGCTTTGGTTAGAGAACTTAAAGACGCTTTTGAACCTCATGGC--- 1093
Qy 151 GluSerAsnArgProArgLeuMetValThrAlaAlaValAlaGlyGlyIleSerAsnIle 170
Db 1092 -----TACTTGTGACTGCTGCAGTATCACCAGGTAAAGACAAATC 1051
Qy 171 GlnAlaGlyTyThrGluIleProGluLeuSerLysTyThrLeuAspPheIleHisValMetThr 190
Db 1050 GACCGAGCTTATGATATCAAGAATTAAGAATAATGTTTCGATTGGATGAATGTCATGACA 991
Qy 191 TyrAspLeuHisGlySerTTPGluGlyTyThrGlyGluAsnSerProLeuTyThrLysTy 210
Db 990 TATGATTACCAAGTGGATGGGAAACTTTTACGCTCAATATGCTCCGTGTATTAACGA 931
Qy 211 ProThrGluThrGlySer---AsnAlaTyThrLeuAsnValAspTyThrValMetAsnTyThr 229
Db 930 CCAGATGAATGATGATGAGTGCACACTTACTTCAATGTCACTACACCATGCACATATTAT 871
Qy 230 LysAsnAsnGlyAlaProAlaGluLysLeuIleValGlyPheProGluTyThrGlyHisThr 249
Db 870 TTGAACAATGGTGGCCACAGAGACAAATTTGGTAAATGGGTGTTCCATTCTATGGCGTGCT 811
Qy 250 PheIleLeuArgAsnProSerAspAsnGlyIleGlyAlaProThrSerGlyAspGlyPro 269
Db 810 TGGAGCATTTGAAGATCGAAGCAAACTCAAACTTGGAGATCCAGCCAAAGGCGATGCGCCC 751
Qy 270 AlaGlyAlaTyThrArgGlnAlaGlyPheTTPAlaTyThrGluIleCysThrPheLeu 289
Db 750 CCAGGTTTCATTTCTGGTGAAGAGTGTCTCTCATATATAGAAATTTGTCAATTGTTT 691
Qy 290 ArgSerGlyAlaThrGluValTTP-----AspAlaSerGlnGluValProTyr 305
Db 690 CAAAAA-----GAAGAATGGCATATCCAAATACGATGAATATATACAAATGCTCCATAT 640
Qy 306 AlaTyThrLysAlaAsnGluTTPLeuGlyTyThrAspAsnIleLysSerPheSerValLysAla 325
Db 639 GGTTCATATGATATAAATCTGGTGGTACGATGATCTGGCCAGTATATCATGCAAGTTG 580
Qy 326 GlnTTPLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTTPAlaIleAspLeuAsp 345
Db 579 GCTTTCCTGAAAGAAATAGGCGTTTCTGGTGCATGTTTGGTCATTGGAAAAATGATGAT 520
Qy 346 PheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSerThrLeuAsnLysAla 365

Db 519 TTCAAAGGT---CACTGCGGA-----CCGAAAAATCCATTGTTGAACAAAGTT 475
 QY 366 -----LeuGlyIleSerThr 370
 Db 474 CATAATATGATTAAATGCGGATGAAGAAGAACTCTTTTGAATGCAATTTGGTTCAGATCA 415
 QY 371 GluGlyCysThrAlaProAspValProSerGluProValThrPro 386
 Db 414 ACGACACCAACTCCACGACGACGACACCCACACCCGACTACAGGCCA 367
 RESULT 13
 US-10-218-743-14
 ; Sequence 14, Application US/10218743
 ; Publication No. US20030096779A1
 ; GENERAL INFORMATION:
 ; APPLICANT: McCall, Catherine A.
 ; APPLICANT: Hunter, Shirley Wu
 ; APPLICANT: Weber, Eric R.
 ; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
 ; AND USES THEREOF
 ; FILE REFERENCE: AL-2-C3
 ; CURRENT APPLICATION NUMBER: US/10/218,743
 ; CURRENT FILING DATE: 2002-08-13
 ; PRIOR APPLICATION NUMBER: US/09/292,225
 ; PRIOR FILING DATE: 1999-04-15
 ; PRIOR APPLICATION NUMBER: 60/098,909
 ; PRIOR FILING DATE: 1998-09-02
 ; PRIOR APPLICATION NUMBER: 60/085,295
 ; PRIOR FILING DATE: 1998-05-13
 ; PRIOR APPLICATION NUMBER: 60/098,565
 ; PRIOR FILING DATE: 1998-04-17
 ; PRIOR APPLICATION NUMBER: 09/062,013
 ; PRIOR FILING DATE: 1998-04-17
 ; NUMBER OF SEQ ID NOS: 49
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 14
 ; LENGTH: 1752
 ; TYPE: DNA
 ; ORGANISM: Dermatophagoides farinae
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(1665)
 ; US-10-218-743-14
 Alignment Scores:
 Pred. No.: 5,35e-75 Length: 1752
 Score: 729.50 Matches: 154
 Percent Similarity: 53.61% Conservative: 69
 Best Local Similarity: 37.02% Mismatches: 144
 Query Match: 29.26% Indels: 49
 DB: 9 Gaps: 11
 US-10-004-219B-9 (1-452) x US-10-218-743-14 (1-1752)
 QY 3 LeuIleCysThrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPheIleThr 22
 Db 103 ATTGTTGTTGTTGTTGAACATGTTCCGATATCATAAA---GTTGATCCATACACTATC 159
 QY 23 AspAspIleAsnProCysLeuGlyCysThrHisLeuIleTyrAlaPheAlaGlyMetGlnAsn 42
 Db 160 GAAGATATTGATCCATTCAAGTGTACACATTTAATGTTATGTTGTTTAAATTTGATGAA 219
 QY 43 AsnGluIleThr-----ileGluTrpAsn 51
 Db 220 TACAATACACAAATTCAGTTTTCGATCTCTTACCAAGATGATAACATAACTATG--- 276
 QY 52 AspValThrLeuTyrIleAlaPheAsnAspLeuIleAsnArgAsnSerIleGlyThr 71
 Db 277 GAAAAACGGTGTATGAACGTTTCAACAACTTCCGATTCAGAAATCCAGAAATTAACCAACC 336
 QY 72 LeuLeuAlaIleGlyCysThrAsnPheGlyThrAlaProPheThrMetValSerThr 91
 Db 337 ATGATTTCACTGGTGGTGGTATGAAGGCTCGGAAAAATATTTCCGATATGCTGCAAT 396

QY 92 SerGlnAsnArgGlnThrPheIleThrSerValIleLysPheLeuArgGlnTyrGlyPhe 111
 Db 397 CCAACATATCGTCAACAAATTCATACAATCAGTTTGGACTTTTTCGAAGAATACAAAGTTC 456
 QY 112 AspGlyLeuAspLeuAspTrpGluTyrProGlySerArg---GlySerProProGlnAsp 130
 Db 457 GACGGTCTAGATTGGATTGGAGTATCTCGATCTCGATTGGTAACCCGAAAAATCGAT 516
 QY 131 LysHisLeuPheThrValLeuValLysGluMetArgGluAlaPheGluGlnGluAlaIle 150
 Db 517 AAACAAACTATTGGCTTTGGTTAGAGAACCTTAAGACGCTTTTGAACCTCATGGC--- 573
 QY 151 GluSerAsnArgProArgLeuMetValThrAlaAlaValAlaGlyGlyIleSerAsnIle 170
 Db 574 -----TACTTGTGACTGCTGCAGTATCACCAGGTAAGACAAAATC 615
 QY 171 GlnAlaGlyTyrGluIleProGluLeuSerLysTyrLeuAspPheIleHisValMetThr 190
 Db 616 GACCGAGCTTATGATATCAAGAATTTGAACAAATTTGTTGATTTGGATGAATGTCATGACA 675
 QY 191 TyrAspLeuHisGlySerTrpGluGlyTyrThrGlyGluAsnSerProLeuTyrLysTyr 210
 Db 676 TATGATTACCACGGTGGATGGGAAACTTTTACGGTCACAATGCTCCGTTGTATAAACA 735
 QY 211 ProThrGluThrGlySer---AsnAlaTyrLeuAsnValAspTyrValMetAsnTyrTrp 229
 Db 736 CCAGATGAACATGATGAGTTCGACACATTTCAATGTCACACTACACCATGACATATTAT 795
 QY 230 LysAsnAsnGlyAlaProAlaGluLysLeuIleValGlyPheProGlyTyrGlyHisThr 249
 Db 796 TTGAACAATGGTGCACCAGACAGACAAATTTGTAATGGGTTCATCTTATGGCCGTGCT 855
 QY 250 PheIleLeuArgAsnProSerAspAsnGlyIleGlyAlaProThrSerGlyAspGlyPro 269
 Db 856 TGGAGCATTTGAAGATCGAAGCAAACTCAAACTTGGAGATCCAGCCAAAGGCGATGTCGCC 915
 QY 270 AlaGlyAlaTyrArgGlnAlaGlyPheTrpAlaTyrTyrGluIleCysThrPheLeu 289
 Db 916 CCAGGTTTCATTCTCGTGAAGAGGTGCTCTCATATATAGAAATTTGTCATATTGTT 975
 QY 290 ArgSerGlyAlaThrGluValTrp-----AspAlaSerGlnGluValProTyr 305
 Db 976 CAAAAA-----GAAGAATGGCATATCCAATACGATGAATATTACAATGCTCCATAT 1026
 QY 306 AlaTyrLysAlaAsnGluTyrLeuGlyTyrAspAsnIleLysSerPheSerValLysAla 325
 Db 1027 GGTTCAAATGATAAAATCTGGGTGCGTTACGATGATCTGGCCAGTATATCATGCAANGTTG 1086
 QY 326 GlnTrpLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTrpAlaIleAspLeuAspAsp 345
 Db 1087 GCTTTCCTGAAAGAATTAGGCGTTTCTGTCATGTTGTTGTCATTCGTAATAATGATCAT 1146
 QY 346 PheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSerThrLeuAsnLysAla 365
 Db 1147 TTCAAAGGT---CACTGCGGA-----CCGAAAAATTCATTTGTTGAACAAAGTT 1191
 QY 366 -----LeuGlyIleSerThr 370
 Db 1192 CATATATGATTAAATGGCGATGAAGAAGAACTCTTTTGAATGCAATTTGGTTCAGATGACA 1251
 QY 371 GluGlyCysThrAlaProAspValProSerGluProValThrPro 386
 Db 1252 ACGACACCAACTCCACGACGACGACCCACCAACCCGACTACACGCCA 1299
 RESULT 14
 US-10-218-743-16/c
 ; Sequence 16, Application US/10218743
 ; Publication No. US20030096779A1
 ; GENERAL INFORMATION:
 ; APPLICANT: McCall, Catherine A.
 ; APPLICANT: Hunter, Shirley Wu
 ; APPLICANT: Weber, Eric R.

; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS

; FILE REFERENCE: AL-2-C3

; CURRENT APPLICATION NUMBER: US/10/218,743

; CURRENT FILING DATE: 2002-08-13

; PRIOR APPLICATION NUMBER: US/09/292,225

; PRIOR FILING DATE: 1999-04-15

; PRIOR APPLICATION NUMBER: 60/098,909

; PRIOR FILING DATE: 1998-09-02

; PRIOR APPLICATION NUMBER: 60/085,295

; PRIOR FILING DATE: 1998-05-13

; PRIOR APPLICATION NUMBER: 60/098,565

; PRIOR FILING DATE: 1998-04-17

; PRIOR APPLICATION NUMBER: 09/062,013

; PRIOR FILING DATE: 1998-04-17

; NUMBER OF SEQ ID NOS: 49

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 16

; LENGTH: 1752

; TYPE: DNA

; ORGANISM: Dermatophagoides farinae

US-10-218-743-16

Alignment Scores:

Pred. No.:	5,35e-75	Length:	1752
Score:	729.50	Matches:	154
Percent Similarity:	53.61%	Conservative:	69
Best Local Similarity:	37.02%	Mismatches:	144
Query Match:	29.28%	Indels:	49
DB:	9	Gaps:	11

US-10-004-219b-9 (1-452) x US-10-218-743-16 (1-1752)

```
QY      3 LeuileCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPheLysPro 22
DB      1650 ATTGTTGTTATGTTGGAAACATGGTCCGTATATCAATAA---GTTGATCCATACACTATC 1594
QY      23 AspAspIleAsnProCysLeuCysThrHisLeuileTyrAlaPheAlaGlyMetGlnAsn 42
DB      1593 GAAGATATTGATCCATTCAAGTGTACACATTAAATGTTGTTGCTTAAATTCATGAA 1534
QY      43 AsnGluileThr-----IleGluTrpAsn 51
DB      1533 TACAATAACACAAATTCAAGTTTTCGATCTTACCAAGATGATAACATAACTCATGG--- 1477
QY      52 AspValThrLeuTyrIleAlaPheAsnAspLeuLysAsnArgAsnSerLysLeuLysThr 71
DB      1476 GAAAAACGTGGTTATGAACGTTTCAACAACTTGCAGATTGAAGAATCCAGAATTAAACACC 1417
QY      72 LeuLeuAlaIleGlyTyrPheThrAsnPheGlyThrAlaProPheThrMetValSerThr 91
DB      1416 ATGATTTCACCTTGGTGGTGGTATGAGGCTCGGAAAAATATTCGATATGGCTCAAAAT 1357
QY      92 SerGlnAsnArgGlnThrPheIleThrSerValIleLysPheLeuArgGlnTyrGlyPhe 111
DB      1356 CCAACATATCGTCAAAATTCACAAATTCAGTTTGGAGCTTTTTCGAAAGATACAAAGTTC 1297
QY      112 AspGlyLeuAspLeuAspTyrGlnTyrProGlySerArg---GlySerProGlnAsp 130
DB      1296 GACGCTGTAGATTTCGATTGGAGAGTATCTCGATTCGATTGGGTAAACCCGAAAAATTCGAT 1237
QY      131 LysHisLeuPheThrValLeuValLysGluMetArgGluAlaPheGluGlnGluAlaIle 150
DB      1236 AAACAAACTATTTCGCTTGGTTAGAGACTTAAAGAGCTTTTGAACCTCATGGC--- 1180
QY      151 GluSerAsnArgProArgLeuMetValThrAlaAlaValAlaGlyLysSerAsnIle 170
DB      1179 -----TACTTGTGCTGTCGAGTATCACCAGTAAACACAAATC 1138
QY      171 GlnAlaGlyTyrGluileProGluLeuSerLysTyrLeuAspPheIleHisValMetThr 190
DB      1137 GACCGAGCTTATGATCAAGAAATTGAACAAATGTTTCGATTGGATGAATGTCATGACA 1078
```

```
QY      191 TyrAspLeuHisGlySerTrpGluGlyTyrThrGlyGluAsnSerProLeuTyrLysTyr 210
DB      1077 TATGATTACCAAGGTGGATGGAAAACTTTTACGGTCACAATGCTCCGTTGATATAACGA 1018
QY      211 ProThrGluThrGlySer---AsnAlaTyrLeuAsnValAspTyrValMetAsnTyrTrp 229
DB      1017 CCAGATGAACACTGATGATGATGCTCACTTACTTCAATGTCAATGACCATACCATGACATATTAT 958
QY      230 LysAsnGlnGlyAlaProAlaGluLeuLeuIleValGlyPheProGluTyrGlyHisThr 249
DB      957 TTGAACAATGGTGCACACAGAGACAAATTGTAATGGGTGTTCATTTATGCGCGTGT 898
QY      250 PheIleLeuArgAsnProSerAspAsnGlyIleGlyAlaProThrSerGlyAspGlyPro 269
DB      897 TGGAGCATTTGAAGATCGAAGCAAACTCAAACTTGAGATCCAGCCAAAGGCATGTCGCC 838
QY      270 AlaGlyAlaTyrThrArgGlnAlaGlyPheTrpAlaTyrTyrGluileCysThrPheLeu 289
DB      837 CCAGGTTTTCATTTCTGCTGTAAGAAGGTGCTCTCATATATAGAATTGTCGCAATTGTTT 778
QY      290 ArgSerGlyAlaThrGluValTrp-----AspAlaSerGlnGluValProTyr 305
DB      777 CAAAAA-----GAAGAATGGCATATCCAAATGAGATGATTAATACATGCTCCATAT 727
QY      306 AlaTyrLysAlaAsnGluTrpLeuGlyTyrAspAsnIleLysSerPheSerValLysAla 325
DB      726 GGTACAAATGATAAAATCTGGTGGTTACCATGATCTGGCCAGTATATCATGCAAGTTG 667
QY      326 GlnTrpLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTrpAlaIleAspLeuAsp 345
DB      666 GCTTTCCTGTAAGAAGAAATTAGGCGTTTCTGCTCATGTTGTCATTCATTCGAAAAATGATGAT 607
QY      346 PheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSerThrLeuAsnLysAla 365
DB      606 TTCAAAGGT---CACTGCGGA-----CCGAAAAATCCATTGTTGAACAAAGTT 562
QY      366 -----LeuGlyIleSerThr 370
DB      561 CATAATATGATTAAATGGCGATGAAAAAGAACTTTTCGAATGCAATTTTCGGTCCCAAGTACA 502
QY      371 GluGlyCysThrAlaProAspValProSerGluProValThrThrPro 386
DB      501 ACGACACCAACTCCCAACGACGACACCCACCAACCCCGACTACAACGCCA 454
RESULT 15
US-10-218-743-40
; Sequence 40, Application US/10218743
; Publication No. US20030096779A1
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
; FILE REFERENCE: AL-2-C3
; CURRENT APPLICATION NUMBER: US/10/218,743
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: US/09/292,225
; PRIOR FILING DATE: 1999-04-15
; PRIOR APPLICATION NUMBER: 60/098,909
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/085,295
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/098,565
; PRIOR FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: 09/062,013
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 40
; LENGTH: 1470
; TYPE: DNA
; ORGANISM: Dermatophagoides farinae
```

FEATURE:
NAME/KEY: CDS
LOCATION: (1) .. (1470)
US-10-218-743-40

Alignment Scores:

Pred. No.: 3,07e-74 Length: 1470
Score: 722.00 Matches: 158
Percent Similarity: 53.96% Conservative: 67
Best Local Similarity: 37.89% Mismatches: 142
Query Match: 28.96% Indels: 50
DB: 9 Gaps: 13

US-10-004-219B-9 (1-452) x US-10-218-743-40 (1-1470)

```
QY 3 LeuileCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPheLysPro 22
DB 46 ATCGTATGTTATGTTGGAACATGGTCGCTTTATCATATAA---GTTGATCCATACACAATT 102
QY 23 AspAspIleAsnProCysLeuCysThrHisLeuileTyrAlaPheAlaGlyMetGlnAsn 42
DB 103 GAAGATATTGATCCCTTCAANTGACTCAATTGATGATGTTGCTAAATCGATGAA 162
QY 43 AsnGluileThrThr-----leGluTrpAsn 51
DB 163 TACAATAACACCATTCACAGTTTGTGATCCATTTCAAGATGATAACCATAACTCATGG--- 219
QY 52 AspValThrLeuTyrLysAlaPheAsnAspLeuLysAsnArgAsnSerLysLeuLysThr 71
DB 220 GAAACACACGGGTATGAACGTTTCAACACTTGAGATTGAAGATTGAAGATTGAACAC 279
QY 72 LeuLeuAlaileGlyLysTrpAsnPheGlyThrAlaProPheThrThrMetValSerThr 91
DB 280 ATGATTTTCAATGGTGGTGGTATGATGAGTTTCAGAAATATTCGATATGCGACCAAT 339
QY 92 SerGlnAsnArgGlnThrPheThrSerValleLysPheLeuArgGlnTyrGlyPhe 111
DB 340 CCAACATATCGTCAGCAATTTGTTCAATCAGTTTGGACTTTTTCAGAGATACAAATTC 399
QY 112 AspGlyLeuAspLeuAspTrpGlnTyrProGlySerArg---GlySerProGlnLasp 130
DB 400 GATGGCTAGATTGGATTGGGAATATCTGGATCACGGTTAGGCAATCTTAAATTCGAT 459
QY 131 LysHisLeuPheThrValLeuValLysGluMetArgGluAlaPheGluGlnGluAlaile 150
DB 460 AAACAAACTATTAACTATTAGTAGAGACTTAAAGAGGCATTTGAACCTTCGCG--- 516
QY 151 GluSerAsnArgProArgLeuMetValThrAlaAlaValAlaGlyLysSerAsnIle 170
DB 517 -----TACTTTGTTGCTGCGCAGTATCACCCGGTAAAGATAAAAT 558
QY 171 GlnAlaGlyTyrGluileProGluLeuSerLysTyrLeuAspPheIleHisValMetThr 190
DB 559 GACGTAGCTTATGAGCTAAAGAAATGAACCAATGTTTCGATTTGGAATGATGTCATGACT 618
QY 191 TyrAspLeuHisGlySerTrpGluGlyTyrThrGlyLysAsnSerProLeuTyrLysTyr 210
DB 619 TATGATTACCATGGCGGATGGGAAATGTTTCGGCCATAATGCTCCGTTGTATAACGA 678
QY 211 ProThrGluThrGlySer---AsnAlaTyrLeuAsnValAspTyrValMetAsnTyrTrp 229
DB 679 CCGGATGAACAGCATGAATTCACACTTACTTCAATGTCACTACACCATGACCATATTAT 738
QY 230 LysAsnAsnGlyValaProAlaGluLysLeuileValGlyPheProGluTyrGlyHisThr 249
DB 739 TTGAACAATGGCGCTACTCGAGACAAACTTGTATGGGTGTTCCATTCATGGTCGTGCT 798
QY 250 PheIleLeuArgAsnProSerAspAsnGlyIleGlyAlaProThrSerGlyAspGlyPro 269
DB 799 TGGAGCATCGAAGATCGAAGCAAGTCAAACTTGGCGATCCGCCCAAGGCATGCTCCT 858
QY 270 AlaGlyAlaTyrThrArgGlnAlaGlyPheThrAlaTyrTyrGluileCysThrPheLeu 289
```

```
DB 859 CCTGGTTTTATTACTGGTGAAGAAGGTGTTCTCTCATATCATGAATTGTGTCAGTTATTC 918
QY 290 ArgSerGlyAlaThrGluValTrp-----AspAlaSerGlnGluValProTyr 305
DB 919 CAGAAA-----GAAGATGGCATATTCAATACGATGATATTACAAATGCTCCATAC 969
QY 306 AlaTyrLysAlaAsnGluTrpLeuGlyTyrAspAsnIleLysSerPheSerValLysAla 325
DB 970 GGATATAATGATAAAATCTGGGTTGGTTACGATGATCTGGCTAGTATATCATGCAAGTTG 1029
QY 326 GlnTrpLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTrpAlaIleAspLeuAspAsp 345
DB 1030 GCCTTTCTCAAGAAGATTGGGCGTCTCTGCGCTTATGATATGTCATTGCGAAACGATGAT 1089
QY 346 PheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSerThrLeuAsnLysAla 365
DB 1090 TTCAAAGGT---CATTCGGACCG---AAATATCCATTG-----TTGACCAAGTT 1134
QY 366 -----LeuGlyIleSerThr 370
DB 1135 CACAATATGATCAATGGTGTGATGAAAAAGAACTCTTACGAATGCTTTTGGGCCCAAGTACA 1194
QY 371 GluGlyCysThrAlaProAspValProSerGluProValThrThrProPro 387
DB 1195 ACCACACCAACA---CCAACACCCCGTCAACTCTTCGACTACCAACCA 1242
```

Search completed: June 30, 2003, 03:57:52

Job time : 203.82 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2003, 20:56:54 / Search time 14.0414 Seconds *
(without alignments)
3094.613 Million cell updates/sec

Title: US-10-004-219B-9

Perfect score: 2493
Sequence: 1 YNLICYFTNWAQYRPLGLSP.....QQHCQAGLVFTSCNCNWP 452

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1334.5	53.5	399	2 S27879	secretory protein
2	1013	40.6	654	2 I38605	oviductal glycopro
3	1010	40.5	537	2 S57197	oviduct-specific g
4	1003	40.2	539	2 I46470	estrogen dependent
5	987.5	39.6	383	2 A49562	cartilage glycopro
6	973.5	39.0	383	2 S51327	heparin-binding gl
7	901	36.1	483	2 A53918	chitinase (EC 3.2.
8	889.5	35.7	617	2 T15408	hypothetical prote
9	863.5	34.6	554	2 A56596	chitinase (EC 3.2.
10	844.5	33.9	525	2 T44445	chitinase (EC 3.2.
11	818.5	32.8	405	2 S61551	breast-regressing
12	810	32.5	504	2 A38221	chitinase (EC 3.2.
13	781.5	31.3	1635	2 T14075	chitinase (EC 3.2.
14	495	19.9	1484	2 T22275	hypothetical prote
15	484	19.4	699	2 A38368	chitinase (EC 3.2.
16	482.5	19.4	1215	2 T43916	chitinase A [impor
17	477.5	19.2	756	2 AB1452	chitinase B homolo
18	475	19.1	599	2 D83764	chitinase BH0916 (
19	472.5	19.0	756	2 AB1088	chitinase B homolo
20	472.5	19.0	831	2 T00323	chitinase (EC 3.2.
21	470.5	18.9	423	2 TQ1975	chitinase (EC 3.2.
22	470.5	18.9	2025	2 T03884	hypothetical prote
23	468.5	18.6	398	2 T04761	chitinase homolog
24	464	18.6	379	2 T04762	chitinase homolog
25	455.5	18.3	424	2 S47133	chitinase (EC 3.2.
26	453.5	18.2	427	2 JC4565	chitinase (EC 3.2.
27	450.5	18.1	424	2 S68121	chitinase I precu
28	442.5	17.7	423	2 S51369	chitinase - fungus
29	418	16.8	378	2 S51591	chitinase (EC 3.2.

30	417.5	16.7	546	2 F84238	chitinase [importe
31	413	16.6	633	2 T24898	hypothetical prote
32	395	15.8	366	2 T04763	chitinase homolog
33	392	15.7	511	2 S61166	probable membrane
34	389	15.6	452	2 JC4038	47K glycoprotein p
35	374	15.0	1054	2 T30933	chitinase (EC 3.2.
36	368	14.8	371	2 T04756	chitinase homolog
37	368	14.8	1051	2 D82428	chitodextrinase VC
38	365.5	14.7	499	2 S04856	chitinase (EC 3.2.
39	364.5	14.6	499	2 S52422	chitinase (EC 3.2.
40	360	14.4	365	2 T04757	chitinase homolog
41	356.5	14.3	332	2 T04754	hypothetical prote
42	355.5	14.3	563	2 S60651	chitinase precursor
43	341.5	13.7	421	2 T04753	hypothetical prote
44	341.5	13.7	1046	2 T30199	chitinase (EC 3.2.
45	333.5	13.4	561	2 A25090	chitinase (EC 3.2.

ALIGNMENTS

RESULT 1

S27879

Secretory protein YM-1 precursor - mouse

C:Species: Mus musculus (house mouse)

C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 22-Jun-1999

C:Accession: S27879

R:Chang, N.C.A.; Liu, C.H.; Chang, A.C.

submitted to the EMBL Data Library, June 1992

A:Description: Molecular characterization of a secretory protein (YM-1) transiently expre

A:Reference number: S27879

A:Accession: S27879

A:Molecule type: mRNA

A:Residues: 1-399 <CHA>

A:Cross-references: EMBL:M94584; NID:G202441; PIDN:AB62394.1; PID:G202442

C:Superfamily: Streptomyces chitinase chi40

F:1-21/Domain: signal sequence #status predicted <SIG>

F:22-399/Product: secretory protein YM-1 #status predicted <MAT>

Query Match 53.5%; Score 1334.5; DB 2; Length 399;
Best Local Similarity 64.5%; Pred.No. 7.6e-92; Mismatches 80; Indels 1; Gaps 1;
Matches 243; Conservative 53;

Qy	1	YNLCYFTNWAQYRPLGLSPKDDINPCLCTHLIYAFAGMNNETITIEWNDVLYKAFN	60
Db	22	YQLMCYVTSWAKDRPIEGSPKGNIDPCLCTHLIYAFAGMNNETITYTHEQDLRDYEALN	81
Qy	61	DLKN-RNSKLTLLAIGQWNGFTAPFTTMVSTSONRQFTITSVKFLROYGPDGLDWE	119
Db	82	GLKDKKNTLKTLLAIGQWNGFTAPFTTMVSTSONRQFTITSVKFLROYGPDGLDWE	141
Qy	120	YPSRGSPQDKHLFTVLKEMREAFQEAIESNRPRLMTAAVAGGISNTOAGYEIPEL	179
Db	142	YPSRGSPQDKHLFTVLKEMREAFQEAIESNRPRLMTAAVAGGISNTOAGYEIPEL	201
Qy	180	SKYLDIFHWMTYDLHSGWEGYTGNSPLYKYPTETGNSNAYLNVDYVMYKNGAPAEKL	239
Db	202	LSLLDYIQWYDLHDPKDGVTGNSPLYKSPYDYGKADNLNVDISIISWKDHGAASEKL	261
Qy	240	IVGPEYGHFTILRNPSNGIGAPTSGDGPAGAYTRQAGFWAYYEICTFLRSGATEVMDA	299
Db	262	IVGPEYGHFTILRNPSNGIGAPTSGDGPAGAYTRQAGFWAYYEICTFLRSGATEVMDA	321
Qy	300	SOEVPYAYKANEWLYGDNIKSFVKAQWLKQNNFGAMWALDDDFGSCDQCKEPLT	359
Db	322	POEVPYAYQNEWLYGDNIVRSFKLKAQWLKQNNFGAMWALDDDFGSCDQCKEPLT	381
Qy	360	STLNKALGISTEGCTAP	376
Db	382	STLKGDLNIHSASCKGP	398

RESULT 2

138605

oviductal glycoprotein - human

C:Species: Homo sapiens (man)

C:Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 29-May-1998

C:Accession: I38605

R:Arrias, E.B.; Verhage, H.G.; Jaffe, R.C.

Biol. Reprod. 51, 685-694, 1994

A:Title: Complementary deoxyribonucleic acid cloning and molecular characterization of a

A:Reference number: I38605; MUID:95119256; PMID:7819450

A:Accession: I38605

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-654 <RES>

A:Cross-references: EMBL:U09550; NID:G529147; PID:G529148

Query Match 40.6%; Score 1013; DB 2; Length 654;

Best Local Similarity 48.8%; Pred. No. 1.4e-67;

Matches 203; Conservative 58; Mismatches 125; Indels 30; Gaps 5;

QY 1 YNLICVFTWQAQVRPGLSGFKPDDINPCLCTHLIYAFAGMNNETITIEWND-VTLTKAF 59

DB 22 HKLVICYFTWHAHRPGPASPILPHDLDPFLCTHLIFAFASNNNNQIVAKOLQDEKILYPEF 81

QY 60 NDLKNRNSKLTLLAIGGNFCTAPFTTVMVSTSONRQTFTITSVIKFLQYGFGLDLWE 119

DB 82 NKLKERNRGLKTLISIGGNFCTVFTTMSLTFNSNRERFVSSVIALLRTHGFDGLDLFPL 141

QY 120 YPGSRGSPQDKHLFTVLVKEMREAFQEAIESNRRLMVTAAVAGGISNIQAGYIPEL 179

DB 142 YPGLRGSPHMDRTWTFLELLQAFKNEAQLTMRPRLLSAAVSGDPHVQKADARLL 201

QY 180 SKYLDPIHMTYDLHGSWEKVTGENSPLYKYPTETGNSNAYLNDVYVMYKKNNGAPAEKL 239

DB 202 GRLLDFISVLSDYDLHGSWEKVTGHNPLFLSGPDPKSSA-----YAMNYWROLGVPEPEKL 256

QY 240 IVGFPEYGHFTILRNPDSNGIGAPTSGDGPAGAYTRQAGFWAYYEICTFLRSGATEVWDA 299

DB 257 IMGIPYGRTHLLKASKNGLOARIGPASPQKTKQEGFLAYFEICFSV-WGAKKHWD 315

QY 300 SOEVPYAYKANWGLGYDNIKSFVSKAQLKQNNFGAMWALDDFTGSCDQKGFPLT 359

DB 316 QYVVPYANKKEWGVGDYDAISFGYKAFKREHFGAMVWTLDDFRGYFCGTGPFPLV 375

QY 360 STLNKALGISTEGCTAPDV-----PSEPVTT-----PBG 388

DB 376 VYLDLILVRAEFSSTSLPQFWLSSAVNSSDTPERLAVTIAWTDTSKILPPGEGAG 431

RESULT 3

S57197

oviduct-specific glycoprotein 95K precursor - bovine (fragment)

C:Species: Bos primigenius taurus (cattle)

C:Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 05-Nov-1999

C:Accession: S57197

R:Sendai, Y.; Abe, H.; Kikuchi, M.; Satoh, T.; Hoshi, H.

Biol. Reprod. 50, 927-934, 1994

A:Title: Purification and molecular cloning of bovine oviduct-specific glycoprotein.

A:Reference number: S57197; MUID:94257768; PMID:8199272

A:Accession: S57197

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-537 <SEN>

A:Cross-references: EMBL:DL6639; NID:G391621; PID:BAA04065.1; PID:dl004583; PID:g391622

C:Keywords: glycoprotein

Query Match 40.5%; Score 1010; DB 2; Length 537;

Best Local Similarity 49.2%; Pred. No. 1.8e-67;

Matches 203; Conservative 53; Mismatches 125; Indels 32; Gaps 6;

QY 1 YNLICVFTWQAQVRPGLSGFKPDDINPCLCTHLIYAFAGMNNETITIEWND-VTLTKAF 59

DB 19 HKLVICYFTWHAHRPGPASPILPHDLDPFLCTHLIFAFASNNNNQIVPKDPQDEKILYPEF 78

QY 60 NDLKNRNSKLTLLAIGGNFCTAPFTTVMVSTSONRQTFTITSVIKFLQYGFGLDLWE 119

DB 79 NKLKERNRGLKTLISIGGNFCTVFTTMSLTFNSNRERFVSSVIALLRTHGFDGLDLFPL 138

QY 120 YPGSRGSPQDKHLFTVLVKEMREAFQEAIESNRRLMVTAAVAGGISNIQAGYIPEL 179

DB 139 YPGLRGSPARDRTWTFVLELLQAFKNEAQLTMRPRLLSAAVSGDPHVQKADARLL 198

QY 180 SKYLDPIHMTYDLHGSWEKVTGENSPLYKYPTETGNSNAYLNDVYVMYKKNNGAPAEKL 239

DB 199 GRLLDFISVLSDYDLHGSWEKVTGHNPLFLSGPDPKSSA-----YAMNYWROLGVPEPEKL 253

QY 240 IVGFPEYGHFTILRNPDSNGIGAPTSGDGPAGAYTRQAGFWAYYEICTFLRSGATEVWDA 299

DB 254 LMGLFTYGRTHLLKASKNGLOARIGPASPQKTKQAGFLAYFEICFVR-AKKRWLN 312

QY 300 SOEVPYAYKANWGLGYDNIKSFVSKAQLKQNNFGAMWALDDFTGSCDQKGFPLT 359

DB 313 QYVVPYANKKEWGVGDYDAISFGYKAFKREHFGAMVWTLDDFRGYFCGTGPFPLV 372

QY 360 STLNKALGISTEGCTAPDV-----PSEPVTT-----PBG 388

DB 373 HTLNLL-VNDEFSSTSPKFWFSTAVNSSRIGPEMPTWTRDLTTGLGILPPG 424

RESULT 4

I46470

estrogen dependent oviduct protein precursor - sheep

C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)

C:Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 05-Nov-1999

C:Accession: I46470

R:DeSouza, M.M.; Murray, M.K.

Endocrinology 136, 2485-2496, 1995

A:Title: An estrogen-dependent secretory protein, which shares identity with chitinases,

and embryo development.

A:Reference number: I46470; MUID:95269691; PMID:7750470

A:Accession: I46470

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-539 <DES>

A:Cross-references: EMBL:U16719; NID:G885600; PID:NAC48471.1; PID:G885601

Query Match 40.2%; Score 1003; DB 2; Length 539;

Best Local Similarity 50.9%; Pred. No. 5.9e-67;

Matches 192; Conservative 62; Mismatches 115; Indels 8; Gaps 4;

QY 1 YNLICVFTWQAQVRPGLSGFKPDDINPCLCTHLIYAFAGMNNETITIE-WNDVTLTKAF 59

DB 22 HKLVICYFTWHAHRPGPASPILPHDLDPFLCTHLIFAFASNNNNQIVPKDPLDEKILYPEF 81

QY 60 NDLKNRNSKLTLLAIGGNFCTAPFTTVMVSTSONRQTFTITSVIKFLQYGFGLDLWE 119

DB 82 NKLKERNRGLKTLISIGGNFCTVFTTMSLTFNSNRERFVSSVIALLRTHGFDGLDLFPL 141

QY 120 YPGSRGSPQDKHLFTVLVKEMREAFQEAIESNRRLMVTAAVAGGISNIQAGYIPEL 179

DB 142 YPGLRGSPARDRTWTFVLELLQAFKNEAQLTMRPRLLSAAVSGDPHVQKADARLL 201

QY 180 SKYLDPIHMTYDLHGSWEKVTGENSPLYKYPTETGNSNAYLNDVYVMYKKNNGAPAEKL 239

DB 202 GRLLDFISVLSDYDLHGSWEKVTGHNPLFLSGPDPKSSA-----YAMNYWROLGVPEPEKL 256

QY 240 IVGFPEYGHFTILRNPDSNGIGAPTSGDGPAGAYTRQAGFWAYYEICTFLRSGATEVWDA 299

DB 257 LMGLFTYGRTHLLKASKNGLOARIGPASPQKTKQAGFLAYFEICFVR-AKKRWLN 315

QY 300 SOEVPYAYKANWGLGYDNIKSFVSKAQLKQNNFGAMWALDDFTGSCDQKGFPLT 359

DB 316 QYVVPYANKKEWGVGDYDAISFGYKAFKREHFGAMVWTLDDFRGNFGCTGPFPLA 375

QY 360 STLNKALGISTEGCTAP 376

DB 376 HTLNLL-VNDEFSSTP 391

```
RESULT 5
A49562
cartilage glycoprotein gp39 precursor - human
N:Alternate names: 39K synovial protein
C:Species: Homo sapiens (man)
C:Date: 23-Mar-1995 #sequence revision 23-Mar-1995 #text_change 22-Jun-1999
R:Accession: A49562; S10677; A33162
R:Hakala, B.E.; White, C.; Recklies, A.D.
J. Biol. Chem. 268, 25803-25810, 1993
A:Title: Human cartilage gp-39, a major secretory product of articular chondrocytes and
A:Reference number: A49562; MUID:94064658; PMID:8245017
A:Accession: A49562
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-383 <HA>
A:Cross-references: GB:M80927; NID:G348911; PIDN:AAAL6074.1; PID:G348912
R:Nyrkos, P.; Golde, E.E.
Biochem. J. 269, 265-268, 1990
A:Title: Human synovial cells secrete a 39 kDa protein similar to a bovine mammary prote
A:Reference number: S10677; MUID:90328983; PMID:2375755
A:Accession: S10677
A:Molecule type: protein
A:Residues: 22-40, 'x', 42-45 <NY2>
C:Superfamily: Streptomyces chitinase chi40
C:Keywords: cartilage; extracellular protein; glycoprotein
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-383/Product: cartilage glycoprotein gp39 #status predicted <WAT>

Query Match 39.6%; Score 987.5; DB 2; Length 383;
Best Local Similarity 50.7%; Pred. No. 5.3e-66;
Matches 186; Conservative 66; Mismatches 106; Indels 9; Gaps 5;

QY 1 YNLICYFTNWAQYRPGLSGSKFPDDINPCLCTHLIYAFAGMNNETIIEWNDVTLKAFN 60
DB 22 YKLVCCYTSWSQYREGDSCFPDADRFCLCTHIIYSFANISNDHIDTWENNDVTLGMLN 81

QY 61 DLKRNRSKLKTLAIGGWNFGTAPFTTMSVTSQNRQTFTITSVKFLROYGFGDGLDWEY 120
DB 82 TLKRNPNLTKLLSVGGWNFGSQRFSKIASNTQSRRTFKSVPPFLRTHGFGDGLDWEY 141

QY 121 PGSRSGPPQDKHLFTVLVKEMREAFQEAIESNRPRMLMTAAVAGGINSIQAGYEIPELS 180
DB 142 PGR-----DKQHTTLIKEMKAEFTKEA-QPKKKQLLSAALSACKVTIDSSYDIKTS 195

QY 181 KYLDFHVMYDHLGWSWEGYTGNSPLYKYPTETGNSNAYLVNDYVMYVWKNNGAPAEKLI 240
DB 196 QHLDIFISIMTYDFHGAWRGTTGHSPLFRQEDASDPRESNTDYAVGYMLRLGAPASKLV 255

QY 241 VGFPEYGHFTILRNPSDNGIGAPTSGDGPAGAYTROAGFWAYYEICTFLRSRGATEVWDAS 300
DB 256 MGIPFGRSFTLAS-SETGVGAPISGPGIPGRFTKEAGLAIYEICDFLR-GATVVRTLG 313

QY 301 QEVPIYAYKANEMWLYDNISFVSKAQWLKQNNFGGAMIWALDIDDDFTGSCDQG-KFPILT 359
DB 314 QQVPYATKGNQWGYDDQESVSKVQYLKDRQLAGAWWALDLDLDDPQSGSCCQDLRFPLT 373

QY 360 STLNKAL 366
DB 374 NAIKDAL 380

RESULT 6
S51327
heparin-binding glycoprotein 38K - pig
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 19-Mar-1997 #sequence revision 19-Mar-1997 #text_change 22-Jun-1999
R:Accession: S51327
R:Shackleton, L.M.; Mann, D.M.; Millis, A.J.T.
submitted to the EMBL Data Library, January 1995
A:Description: Identification of a 38kDa heparin-binding glycoprotein (gp38k) in differe
A:Reference number: S51327
```

```
A:Accession: S51327
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-383 <SHA>
A:Cross-references: EMBL:Z47803; NID:G634097; PIDN:CAA87764.1; PID:G634098
C:Superfamily: Streptomyces chitinase chi40

Query Match 39.0%; Score 973.5; DB 2; Length 383;
Best Local Similarity 50.4%; Pred. No. 5.9e-65;
Matches 185; Conservative 63; Mismatches 110; Indels 9; Gaps 5;

QY 1 YNLICYFTNWAQYRPGLSGSKFPDDINPCLCTHLIYAFAGMNNETIIEWNDVTLKAFN 60
DB 22 YKLVCCYTSWSQYREGDSCFPDADRFCLCTHIIYSFANISNDHIDTWENNDVTLGMLN 81

QY 61 DLKRNRSKLKTLAIGGWNFGTAPFTTMSVTSQNRQTFTITSVKFLROYGFGDGLDWEY 120
DB 82 TLKRNPNLTKLLSVGGWNFGSQRFSKIASNTQSRRTFKSVPPFLRTHGFGDGLDWEY 141

QY 121 PGSRSGPPQDKHLFTVLVKEMREAFQEAIESNRPRMLMTAAVAGGINSIQAGYEIPELS 180
DB 142 PGR-----DKRHLTTLVKEMKAEFVREALPTE-RLLSGAVSAGKVAIDRGYDIAQIS 195

QY 181 KYLDFHVMYDHLGWSWEGYTGNSPLYKYPTETGNSNAYLVNDYVMYVWKNNGAPAEKLI 240
DB 196 QHLDIFISIMTYDFHGAWRGTTGHSPLFRQEDASDPRESNTDYAVGYMLRLGAPASKLV 255

QY 241 VGFPEYGHFTILRNPSDNGIGAPTSGDGPAGAYTROAGFWAYYEICTFLRSRGATEVWDAS 300
DB 256 MGIPFGRSFTLAS-SKTDVGAPASGPGIPGRFTKEGILAIYEICDFLR-GATVVRTLG 313

QY 301 QEVPIYAYKANEMWLYDNISFVSKAQWLKQNNFGGAMIWALDIDDDFTGSCDQG-KFPILT 359
DB 314 QQVPYATKGNQWGYDDQESVSKVQYLKDRQLAGAWWALDLDLDDPQSGSCCQDLRFPLT 373

QY 360 STLNKAL 366
DB 374 SAIKDVL 380

RESULT 7
A53918
chitinase (EC 3.2.1.14) precursor - braconid wasp (Chelonus sp.)
C:Species: Chelonus sp.
C:Date: 28-Jul-1995 #sequence revision 28-Jul-1995 #text_change 21-Jul-2000
C:Accession: A53918
R:Krishnan, A.; Nair, P.N.; Jones, D.
J. Biol. Chem. 269, 20971-20976, 1994
A:Title: Isolation, cloning, and characterization of new chitinase stored in active form
A:Reference number: A53918; MUID:94342256; PMID:8063715
A:Accession: A53918
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-483 <KRI>
A:Cross-references: GB:U10422; NID:G533504; PIDN:AAAG1639.1; PID:G533505
C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 36.1%; Score 901; DB 2; Length 483;
Best Local Similarity 38.6%; Pred. No. 2e-59;
Matches 184; Conservative 77; Mismatches 170; Indels 46; Gaps 10;

QY 3 LICYFTNWAQYRPGLSGSKFPDDINPCLCTHLIYAFAGMNNETIIE-WND----VTLK 57
DB 24 VVCYFGASVYRQGNKGFIDIDTFLCTHLIYFVGNGKDVKVLDPHSDLPGLNDGFG 83

QY 58 AFNDLKNRNSKLKTLAIGGWNFGTAPFTTMSVTSQNRQTFTITSVKFLROYGFGDGLD 117
DB 84 KFTSLRKNKPSVKIMVAVGWNAGSVPSQMASDQATREAFQNVVYKFLQYQYQDFG 143

QY 118 WEYPSRSGSPPODKHLFTVLVKEMREAFQEAIESNRPRMLMTAAVAGGINSIQAGYE 177
DB 144 WEYPAQRGGSPADVKNMVKLCKALKKAFVQH-----DYILSAAVAAPETSASKSYDIA 196
```


A:Residues: 1-699 <WAT>
A:Cross-references: GB:M57601; GB:J05599; NID:gl066341; PIDN:AAA81528.1; PID:gl42688
C:Superfamily: fibronectin type III repeat homology
C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 19.4%; Score 484; DB 2; Length 699;
Beat Local Similarity 27.4%; Pred. No. 4.1e-28;
Matches 130; Conservative 80; Mismatches 156; Indels 108; Gaps 18;

```
QY 1 YNLICYFTNWAQYRGLGSKFPDDINPCLCTHLIYAFAGM----- 40
Db 44 YKIVGYFSSWAAYG---RNYNVADIDPKVTHINYAFADICWNGIHGNDPDSGPNPVTT 100

QY 41 -QNNBITTIE-----WNVTLYKAF-----NDLKRNNSKLKTL 72
Db 101 CQNEKSQIINVENGTIVLGDPIWDG--KTFAGDTWDQPIAGNINQNLNKLKQTNENLAKTI 158

QY 73 LAIGQWNGTAPFTTWSTSQNRQFTITSVKFLROYGFDGLDLWEYFGSRG-----SP 127
Db 159 ISVGGTWSNR-FSDVAATAATREVFANSADVFLRKYNFDGVDLDWEYFVSGGLDGNKR 217

QY 128 PQDKHLFTVLKEMEREAEQEAIESNRRLMYTAAGGISNIQAGYEIPELSKYLDRTH 187
Db 218 PEDKQNYTLKLSKIREKLDAAAGVDGKKYLLTIA--SGASATYAANTELAIAIWDWIN 275

QY 188 VMTYDLHGSWEGYTGENSEPLYKYPTET-----GSNAYLNVDYVMNYWKNG 233
Db 276 IMTYDFNGAWQKISAHNAPLNVDPAASAAGVDPDANTFNVAAGAQGHDA-----G 325

QY 234 APAEKLIVGFPEYHTFILRNPDSNGIGAPTSQDGDPAGAYTRQAGWAYYEI-CTFL-RS 291
Db 326 VPAAKLVLGVPFYGRGWDGCAQAGNGQYCTCTGSSVG--TWEAGSFDFYDLEANNYINKN 383

QY 292 GATEVWDASQEVPIYAYKAN--EWLGYDNIKSFVKAQWLKQNNFGGAMIWALDLDFTGS 349
Db 384 GYTRVWNTAKVPYLYNASNKRPFISYDDAESVGYKTAYIKSKGLGAMFW-----ELSG- 437

QY 350 FCDQKFLTLSTLNKALGISTEGCTAPDVPSPPVTPPGSGSGSGSGSGSGGS 403
Db 438 --DRNK-----TLQNLKADLPTGGTVP-----PVDTTAPSVFGNARSTGVTANS 480
```

Search completed: June 29, 2003, 21:02:40
Job time : 16.0414 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2003, 20:55:24 ; Search time 25.4083 Seconds
(without alignments)
3665.470 Million cell updates/sec

Title: US-10-004-219B-9

Perfect score: 2493

Sequence: 1 YNLICYFTNQAQYRPLGSLF.....QQHCQAGLVFTSCNCNWP 452

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 21.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_virus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Description
1	2493	100.0	Q99ph2 mus musculus
2	2488	99.8	Q9jln1 mus musculus
3	2488	99.8	Q9d803 mus musculus
4	2089.5	83.8	Q9b2p6 homo sapien
5	2000	80.2	Q91xa9 mus musculus
6	1991.5	79.9	Q95m17 bos taurus
7	1663.5	66.7	Q9uly4 homo sapien
8	1406.5	56.4	Q9uly3 homo sapien
9	1403	56.3	Q9d7w6 mus musculus
10	1387	55.6	O35744 mus musculus
11	1379	55.3	P70201 mus musculus
12	1376	55.2	Q91z98 mus musculus
13	1373	55.1	Q8vng1 mus musculus
14	1371	55.0	Q90w34 bufo japoni
15	1368	54.9	Q8vh43 mus musculus
16	1301.5	52.2	Q13231 homo sapien

17	1146	46.0	387	4	Q9H3V8
18	1009.5	40.5	624	6	O19118 macaca mula
19	1008	40.4	696	5	Q9VZV2
20	995	39.9	1013	5	Q960M0
21	985.5	39.5	383	4	Q96H17
22	976.5	39.2	457	6	Q95L83
23	973.5	39.0	383	6	Q29411
24	972.5	39.0	383	6	Q8SPQ0
25	966.5	38.8	352	11	Q9WTV1
26	951	38.1	260	11	O61201 mus musculus
27	944.5	37.9	390	4	Q96F97
28	922	37.0	396	11	Q9D7Q1
29	901	36.1	483	5	Q23737
30	900.5	36.1	462	5	Q8SYH0
31	892.5	35.8	381	11	Q99J84
32	879.5	35.3	460	5	Q9W2M7
33	879	35.3	332	6	O18949
34	860.5	34.5	544	5	Q9GQC4
35	859.5	34.5	566	5	Q8WR52
36	858	34.4	553	5	P91731
37	857.5	34.4	543	5	Q9GV05
38	857.5	34.4	543	5	Q9GR93
39	857.5	34.4	565	5	P90710
40	851	34.1	4498	5	Q9W223
41	849.5	34.1	552	5	Q9GV44
42	844.5	33.9	525	5	O44079
43	841.5	33.8	565	5	Q9GPG9
44	838.5	33.6	460	5	Q8WS95
45	826.5	33.2	520	5	Q17100

ALIGNMENTS

RESULT 1

Q99PH2 ID: Q99PH2 PRELIMINARY; PRT; 473 AA.
AC Q99PH2; DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DE 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Acidic mammalian chitinase (EC 3.2.1.14).
GN CHIA.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C; TISSUE=LUNG;
RX MEDLINE=21125893; PubMed=11085997;
RA Boot R.G., Blommaert E.F.C., Swart E., Chauharali-Van Der Vlugt K.,
RA Bijl N., Moe C., Place A., Aerts J.M.F.G.;
RT "Identification of a Novel Acidic Mammalian Chitinase Distinct from
RT Chitotriosidase.";
RL J. Biol. Chem. 276:6770-6778(2001).
CC - SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
CC HYDROLASES).
DR EMBL; AF290003; AAG60018.1; -.
DR MGD; MGI:1932052; Chia.
DR InterPro; IPR001579; Chitinase 18/2.
DR InterPro; IPR002557; Chitin bind PerA.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF01607; CEM_14; 1.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR SMART; SM00494; ChtBD2; 1.
DR PROSITE; PS01095; CHITINASE_18; 1.
KW Glycosidase; Hydrolase
SQ SEQUENCE. 473 AA; 51977 MW; 389D87557BEC0784 CRC64;
Query Match 100.0%; Score 2493; DB 11; Length 473;
Best Local Similarity 100.0%; Pred. No. 4.8e-171;

```
Matches 452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 YNLICFTNWAQVRGLGSKFPDDINPCLCTHLIYAFAGMNNNEITTEWNDVTLKAFN 60
DB 22 YNLICFTNWAQVRGLGSKFPDDINPCLCTHLIYAFAGMNNNEITTEWNDVTLKAFN 81
QY 61 DLKRNRSKLTLLAIGGNWFGTAPFTTWSVTSQNRQTFITSVIKFLRQYGFGLDLDEY 120
DB 82 DLKRNRSKLTLLAIGGNWFGTAPFTTWSVTSQNRQTFITSVIKFLRQYGFGLDLDEY 141
QY 121 PGSRGSPQDKHLFTVLVKEMREAFQEAIESNRPRLMTAAVAGGINSIQAGYEIPELS 180
DB 142 PGSRGSPQDKHLFTVLVKEMREAFQEAIESNRPRLMTAAVAGGINSIQAGYEIPELS 201
QY 181 KYLDFTHTVMTYDLHGSWEGYTGNSPLKYPTETGSNAYLVNDVYNNYKNGAPAEKLI 240
DB 202 KYLDFTHTVMTYDLHGSWEGYTGNSPLKYPTETGSNAYLVNDVYNNYKNGAPAEKLI 261
QY 241 VGPEYGHFTILRNPSDNGIGAPTSQDGPAGAYTRQAGFWAYVEICTFLRSGATEVWDAS 300
DB 262 VGPEYGHFTILRNPSDNGIGAPTSQDGPAGAYTRQAGFWAYVEICTFLRSGATEVWDAS 321
QY 301 QVVPYAYKANWELGYDNIKSFVKAQWLKQNNFGGAMIWAIDLDDFTGSCDQKPEPLTS 360
DB 322 QVVPYAYKANWELGYDNIKSFVKAQWLKQNNFGGAMIWAIDLDDFTGSCDQKPEPLTS 381
QY 361 TLNKGALISTEGCTAPDVPSEPTTPPGSGSGSGSGSGSGSGGFCADKADGLYPVADDR 420
DB 382 TLNKGALISTEGCTAPDVPSEPTTPPGSGSGSGSGSGSGSGGFCADKADGLYPVADDR 441
QY 421 NAFWQINGITYOQHCOAGLVFDTSCNCCNWP 452
DB 442 NAFWQINGITYOQHCOAGLVFDTSCNCCNWP 473

RESULT 2
Q9JLN1 PRELIMINARY; PRT; 472 AA.
ID AC Q9JLN1
AC Q9JLN1
DT 01-OCT-2000 (T-EMBLrel. 15, Created)
DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
DE Putative chitinase precursor (Fragment).
GN CHIA OR YNL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC147666; TISSUE=SKIN;
RA Price P.A., Harris S.C., Williamson M.K.;
RT "YNL, A Putative Mouse Chitinase.";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
CC HYDROLASES).
DR EMBL; AF154571; AAF31644.1; -.
DR MGD; MG1:132052; Chitinase_18/2.
DR InterPro; IPR001579; Chitinase_18/2.
DR InterPro; IPR002557; Chitin_bind_Pera.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF01607; CBM 14; 1.
DR Pfam; PF0704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR SMART; SM00494; ChitB2; 1.
DR PROSITE; PS01095; CHITINASE_18; 1.
KW Glycosidase; Hydrolase; Signal.
FT SIGNAL 1 1
FT CHAIN 21 472 PUTATIVE CHITINASE.
SQ SEQUENCE 472 AA; 51872 MW; FFF59088512C8A7F0 CRC64;
Query Match 99.8%; Score 2488; DB 11; Length 472;
```

```
Best Local Similarity 99.8%; Pred. No. 1.1e-170;
Matches 451; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 YNLICFTNWAQVRGLGSKFPDDINPCLCTHLIYAFAGMNNNEITTEWNDVTLKAFN 60
DB 21 YNLICFTNWAQVRGLGSKFPDDINPCLCTHLIYAFAGMNNNEITTEWNDVTLKAFN 80
QY 61 DLKRNRSKLTLLAIGGNWFGTAPFTTWSVTSQNRQTFITSVIKFLRQYGFGLDLDEY 120
DB 81 DLKRNRSKLTLLAIGGNWFGTAPFTTWSVTSQNRQTFITSVIKFLRQYGFGLDLDEY 140
QY 121 PGSRGSPQDKHLFTVLVKEMREAFQEAIESNRPRLMTAAVAGGINSIQAGYEIPELS 180
DB 141 PGSRGSPQDKHLFTVLVKEMREAFQEAIESNRPRLMTAAVAGGINSIQAGYEIPELS 200
QY 181 KYLDFTHTVMTYDLHGSWEGYTGNSPLKYPTETGSNAYLVNDVYNNYKNGAPAEKLI 240
DB 201 KYLDFTHTVMTYDLHGSWEGYTGNSPLKYPTETGSNAYLVNDVYNNYKNGAPAEKLI 260
QY 241 VGPEYGHFTILRNPSDNGIGAPTSQDGPAGAYTRQAGFWAYVEICTFLRSGATEVWDAS 300
DB 261 VGPEYGHFTILRNPSDNGIGAPTSQDGPAGAYTRQAGFWAYVEICTFLRSGATEVWDAS 320
QY 301 QVVPYAYKANWELGYDNIKSFVKAQWLKQNNFGGAMIWAIDLDDFTGSCDQKPEPLTS 360
DB 321 QVVPYAYKANWELGYDNIKSFVKAQWLKQNNFGGAMIWAIDLDDFTGSCDQKPEPLTS 380
QY 361 TLNKGALISTEGCTAPDVPSEPTTPPGSGSGSGSGSGSGGFCADKADGLYPVADDR 420
DB 381 TLNKGALISTEGCTAPDVPSEPTTPPGSGSGSGSGSGSGGFCADKADGLYPVADDR 440
QY 421 NAFWQINGITYOQHCOAGLVFDTSCNCCNWP 452
DB 441 NAFWQINGITYOQHCOAGLVFDTSCNCCNWP 472

RESULT 3
Q9D803 PRELIMINARY; PRT; 473 AA.
ID AC Q9D803
AC Q9D803
DT 01-JUN-2001 (T-EMBLrel. 17, Created)
DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
DE 220003E03Rik protein.
GN CHIA OR 220003E03RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=STOMACH;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King S., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Stauble F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Hayashizaki Y.;
RA Hayashizaki A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
```



```
DR Pfam; PF01607; CBM_14; 1.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR PROSITE; PS01095; CHITINASE_18; UNKNOWN 1.
SQ SEQUENCE 365 AA; 39968 MW; C900BE0C4416F1DD CRC64;

Query Match
Best Local Similarity 99.7%; Score 2000; DB 11; Length 365;
Matches 364; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 88 MVSTSONQRTFITSVVKFLROYGDFGLDLDWEYVPGSRGSPQDKHLFTVLVKEMREAFQ 147
DB 1 MVSTSONQRTFITSVVKFLROYGDFGLDLDWEYVPGSRGSPQDKHLFTVLVKEMREAFQ 60

QY 148 EATESNRPRMLMTAAVAGGINSIQAGVEIPELSKYLDLHVMTYDLHGSWEGYTGNSPL 207
DB 61 EATESNRPRMLMTAAVAGGINSIQAGVEIPELSKYLDLHVMTYDLHGSWEGYTGNSPL 120

QY 208 YKPTTGTGSNAYLNVDYVMNYKNGNGAPAEKLIIVGPEYGHFTILRNPSDNGIGAPTSGD 267
DB 121 YKPTTGTGSNAYLNVDYVMNYKNGNGAPAEKLIIVGPEYGHFTILRNPSDNGIGAPTSGD 180

QY 268 GPAGATYRQAGFWAYYEICTFLRSQATEVMDASQEVYAYKANEMWIGYDNIKSFVKAQW 327
DB 181 GPAGPYTRQAGFWAYYEICTFLRSQATEVMDASQEVYAYKANEMWIGYDNIKSFVKAQW 240

QY 328 LKONNFGGAMIWALDLDFTGSCDQKPLTSTLNKALGISTEGCTAPDVPSEPTTTP 387
DB 241 LKONNFGGAMIWALDLDFTGSCDQKPLTSTLNKALGISTEGCTAPDVPSEPTTTP 300

QY 388 GSGSGSGSGSGSGSGGFCADKADGLYPVADDRNAPFQINGITYQOHCQAGLVFTSCN 447
DB 301 GSGSGSGSGSGSGSGGFCADKADGLYPVADDRNAPFQINGITYQOHCQAGLVFTSCN 360

QY 448 CCNWP 452
DB 361 CCNWP 365

RESULT 6
Q95M17 PRELIMINARY; PRT; 472 AA.
AC Q95M17;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Chitin binding protein b04.
GN CBP B04.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE=21475601; PubMed=11591385;
RA Suzuki M., Morimatsu M., Yanashita T., Iwanaga T., Syuto B.;
RT "A novel serum chitinase that is expressed in bovine liver.";
RL FEBS Lett. 506:127-130(2001).
DR EMBL; AB051629; BAB71805.1; -.
DR InterPro; IPR001579; Chitinase_18/2.
DR InterPro; IPR002557; Chitin_bind_Pera.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF01607; CBM_14; 1.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR PROSITE; PS01095; CHITINASE_18; UNKNOWN 1.
SQ SEQUENCE 472 AA; 52129 MW; 7A4A600E8DA04B1E CRC64;

Query Match
Best Local Similarity 79.9%; Score 1991.5; DB 6; Length 472;
Matches 348; Conservative 45; Mismatches 57; Indels 1; Gaps 1;

QY 88 MVSTSONQRTFITSVVKFLROYGDFGLDLDWEYVPGSRGSPQDKHLFTVLVKEMREAFQ 147
DB 1 MVSTPENQRTFITSVVKFLROYGDFGLDLDWEYVPGSRGSPQDKHLFTVLQVEMREAFQ 60
```

```
QY 1 YNLICVFTNWAQYRPLGSGFKPDDINPCLCTHLIYAFAGMONEITTTIENWVTLTKAFN 60
DB 22 YQVLCVFNWAQYRPLGSGFKPDDINPCLCTHLIYAFAGMSEITTTIENWVTLTKAFN 81

QY 61 DLKRNRSKLTLLAIGGNWFGTAPFTTVMVSTSONQRTFITSVVKFLROYGDFGLDLDWEY 120
DB 82 DLKKNSQLKILLAIGGNWFGTAPFTTVMVSTSONQRTFITSVVKFLROYGDFGLDLDWEY 141

QY 121 PGRSGSPQDKHLFTVLVKEMREAFQEAIESNRPRMLMTAAVAGGINSIQAGYIPELS 180
DB 142 PGRSGSPQDKHLFTVLVKEMREAFQEAIESNRPRMLMTAAVAGGINSIQAGYIPELS 201

QY 181 KYLDLHVMTYDLHGSWEGYTGNSPLKYPTTGTGSNAYLNVDYVMNYKNGNGAPAEKLI 240
DB 202 QYLDLHVMTYDLHGSWEGYTGNSPLKYPTTGTGSNAYLNVDYVMNYKNGNGAPAEKLI 261

QY 241 VGFPEYGHFTILRNPSDNGIGAPTSGDGPAGATYRQAGFWAYYEICTFLRSQATEVMDAS 300
DB 262 IGFPAYGHNFLIRDASNNIGAPTSGDGPAGATYRQAGFWAYYEICTFLRSQATEVMDAS 321

QY 301 QVPYAYKANEMWIGYDNIKSFVKAQWLNKONNFGGAMIWALDLDFTGSCDQKPLT 360
DB 322 QNVYAYKGTETWGVYDNVNSPRIKAQWLNKONNFGGAMIWALDLDFTGSCDQKPLT 381

QY 361 TLNKGALGISTEGCTAPDVPSEPTTTPGSGSGSGSGSGGFCADKADGLYPVADDR 420
DB 382 TLKDALGLKATCATNSTQSSP-NSPNEGSGNGKSSSESGRGYCACGADGLYPVADDR 440

QY 421 NAFWOCINGITYQOHCQAGLVFTSCNCCNW 451
DB 441 NAFWNCVNGITYQKONCLTGLVFTSCHKCNW 471

RESULT 7
Q9ULY4 PRELIMINARY; PRT; 368 AA.
AC Q9ULY4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Novel member of chitinase family.
GN TSA1902-L.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20018184; PubMed=10548734;
RA Saito A., Ozaki K., Fujiwara T., Nakamura Y., Tanigami A.;
RT "Isolation and mapping of a human lung-specific gene, TSA 1902, encoding a novel chitinase family member.";
RL Gene 239:325-331(1999).
DR EMBL; AB025008; BAA86980.1; -.
DR InterPro; IPR001579; Chitinase_18/2.
DR InterPro; IPR002557; Chitin_bind_Pera.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF01607; CBM_14; 1.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR SMART; SM00494; ChitBD2; 1.
DR PROSITE; PS01095; CHITINASE_18; UNKNOWN 1.
SQ SEQUENCE 368 AA; 40082 MW; 10FB970C79E19254 CRC64;

Query Match
Best Local Similarity 79.8%; Score 1663.5; DB 4; Length 368;
Matches 293; Conservative 36; Mismatches 35; Indels 3; Gaps 1;

QY 88 MVSTSONQRTFITSVVKFLROYGDFGLDLDWEYVPGSRGSPQDKHLFTVLVKEMREAFQ 147
DB 1 MVSTPENQRTFITSVVKFLROYGDFGLDLDWEYVPGSRGSPQDKHLFTVLQVEMREAFQ 60
```

```
QY 148 EAIESNRLMLMTAAVAGGISNIQAGYEIPELSKYLDHFHVMTYDLHGSWEGYTGENSPL 207
  ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 EAKQINKPRLMTAAVAGGISNIQSGYEIPQLSYLDYHVMYDLHGSWEGYTGENSPL 120
QY 208 YKYPTEGSGNAYLVNYYWYKNGAPAEKLIYVGPPEYGHFTFLRNSDNGIGAPTSGD 267
  ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 YKYPTEGSGNAYLVNYYWYKNGAPAEKLIYVGPPEYGHFTFLRNSDNGIGAPTSGA 180
QY 268 GPAGYATROAGFWAYYEICTFLRSGATEVWDASQEVYAYKANEMWLDYDNKISFVSQAQW 327
  ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 181 GPAGYAKESGIWAYYEICTFLKNGATQWDAPQEVYAYQGVNVMYGVNDFDKAQW 240
QY 328 LKQNFPGGAMIWAILDDFTGSCDQKRPFLSTLTKALGISTEGCTAPDVPSEPVTTTP 387
  ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 241 LKHNFGGAMWAILDDFTGTCNQKRPFLSTLTKALGLQASCTAPAPQIEPITRAP 300
QY 388 ---GSGSGSGSGSGSGGFCADKADGLYPVADDRNAPWQINGITTYOQHCOAGLVFDT 444
  ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 301 SSGSGSGSGSGSGSGGFCAGRANGLYPVANNRNPFWHCVCVGTYYQNCQAGLVFDT 360
QY 445 SCNCCNW 451
  ||:|||||
Db 361 SCDCCNW 367
  ||:|||||

RESULT 8
Q9ULY3 PRELIMINARY; PRT; 315 AA.
AC Q9ULY3;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Novel member of chitinase family.
CN TSA1902-S.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20018184; PubMed=10548734;
RA Saico A., Ozaki K., Fujiwara T., Nakamura Y., Tanigami A.;
RT "Isolation and mapping of a human lung-specific gene, TSA 1902,
RT encoding a novel chitinase family member.";
RL Gene 239:325-331(1999).
DR EMBL; AB025009; BAA86981.1; -
DR InterPro; IPR002557; Chitin_bind_Pera.
DR Pfam; PF01607; CBM_14; 1.
DR ProDom; PD00704; Glyco_hydro_18; 1.
DR SMART; SM00494; ChtBD2; 1.
SQ SEQUENCE 315 AA; 33849 MW; 8D08DB82C1268F0 CRC64;

Query Match 56.4%; Score 1406.5; DB 4; Length 315;
Best Local Similarity 78.0%; Pred. No. 3,1e-93;
Matches 245; Conservative 34; Mismatches 32; Indels 3; Gaps 1;

QY 141 MREAPEQEAIESNRLMLMTAAVAGGISNIQAGYEIPELSKYLDHFHVMTYDLHGSWEGY 200
  ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MREAPEQEAQINKPRLMTAAVAGGISNIQSGYEIPQLSYLDYHVMYDLHGSWEGY 60
QY 201 TCENSPLYKYPTETGSGNAYLVNYYWYKNGAPAEKLIYVGPPEYGHFTFLRNSDNGI 260
  ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 TGENSEPLKYPTETGSGNAYLVNYYWYKNGAPAEKLIYVGPPEYGHFTFLRNSDNGI 120
QY 261 GAPTSGDGPAGYATROAGFWAYYEICTFLRSGATEVWDASQEVYAYKANEMWLDYDNKIS 320
  ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 GAPTSGAGPAGYAKESGIWAYYEICTFLKNGATQWDAPQEVYAYQGVNVMYGVNDFDK 180
QY 321 FSVKAQWLKQNFPGGAMIWAILDDFTGSCDQKRPFLSTLTKALGISTEGCTAPDVP 380
  ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 181 FDIKAQWLKHNKFGGAMIWAILDDFTGTCNQKRPFLSTLTKALGLQASCTAPAP 240
```

```
QY 381 EPTVTPP---GSGSGSGSGSGSGGFCADKADGLYPVADDRNAPWQINGITTYOQHCO 437
  ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 241 EPTAAPSGSGSGSGSGSGGFCAGRANGLYPVANNRNPFWHCVCVGTYYQNCQ 300
QY 438 AGLVFDTSNCNCNW 451
  ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 301 AGLVFDTSDCCNW 314

RESULT 9
Q9D7W6 PRELIMINARY; PRT; 262 AA.
AC Q9D7W6;
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-WAR-2002 (TRENBLrel. 20, Last annotation update)
DE 2200003E03Rik protein.
GN CHIA OR 2200003E03RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=STOMACH;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yananaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Glasi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Mateau Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynehaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -!- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
CC HYDROLASES).
DR EMBL; AK008757; BAB25878.1; -
DR MGD; MGI:1932052; Chia.
DR InterPro; IPR001579; Chitinase_18/2.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR PROSITE; PS01095; CHITINASE_18; 1.
KW Glycosidase; Hydrolase.
SQ SEQUENCE 262 AA; 29458 MW; 954AA0F0ELC9851D CRC64;

Query Match 56.3%; Score 1403; DB 11; Length 262;
Best Local Similarity 99.2%; Pred. No. 4.4e-93;
Matches 260; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 40 MQNNEITTIENVDVLYKAFNDLKRNKSLKTLAIGGNFGTAPFTTMVSTSQNRQTFI 99
  ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MQNNEITTIENVDVLYKAFNDLKRNKSLKTLAIGGNFGTAPFTTMVSTSQNRQTFI 60
QY 100 TSVIKFLQYFGDGLDWEYFGSGSPQDKHLFTVLVKEMREAFEOEATESNPRLMV 159
  ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 TSVIKFLQYFGDGLDWEYFGSGSPQDKHLFTVLVKEMREAFEOEATESNPRLMV 120
QY 160 TAAVAGGISNIQAGYEIPELSKYLDHFHVMTYDLHGSWEGYTGENSPLYKYPTETG 219
  ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```


OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=BALB/C;
RY PubMed=11533626;
RA Webb D.C., McKenzie A.N.J., Foster P.S.;
"Expression of the Ym2 Lectin-binding Protein Is Dependent on
RT Interleukin (IL)-4 and IL-13 Signal Transduction. IDENTIFICATION OF A
RT NOVEL ALLERGY-ASSOCIATED PROTEIN";
RL J. Biol. Chem. 276:41969-41976(2001).
DR EMBL: AV049765; AAL03953.1; -;
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
KW Signal.
FT SIGNAL.
SQ SEQUENCE 1 21 POTENTIAL
SEQUENCE 398 AA; 44515 MW; 79A74CA5CDA03850 CRC64;
Query Match 55.2%; Score 1376; DB 11; Length 398;
Best Local Similarity 65.2%; Pred. No. 6.7e-91;
Matches 245; Conservative 61; Mismatches 70; Indels 0; Gaps 0;
QY 1 YNLICFTWTAQYRPGLSFKFDDINPCLCTHLIYAFAGMKNNEITTIWINDVTLKAPN 60
DB 22 YQLMCYITSWAKDRPTEGSGFKGNIDPCCLCTHLIYAFAGMKNNEITYLSEQDLRDYEALN 81
QY 61 DLKRNRSKLKTLAIGGWNFGTAPFTTWTSTONRQTFTSVIKFLQYGFGLDLDMWEY 120
DB 82 GLKDRNTELTLLAIGGWNFGTAPFTTWTSTONRQTFTSVIKFLQYGFGLDLDMWEY 141
QY 121 PGSRSGPPQDKHLFTVLKEMREAFQEAIESNRPRLMVTAAGVAGISNIQAGYIPELS 180
DB 142 PGSRSGPPKDKHLFSLVQEMRKAFEEESTLNHPRLLLTSTGAGFIDVIGSYKIPELS 201
QY 181 KYLDFIHWMTYDLHSGWEGYTGNSPLYKYPTETGSNAYLVNDVYVMYKNGNAPAEKLI 240
DB 202 QSLDYIQVMTYDLHDPKNGYTGNSPLYKSPYDICKSADLNVDSTIYWKDHGAASEKLI 261
QY 241 VGFPEYGHFTILRNPDSNGIGAPTSGDPAGAYTRQAGFWAYYEICTFLRSGATEVMDAS 300
DB 262 VGFPAVGHFTILSDPSKNGIGDPTVSAGPPGKYTNEQGLLAYFEICTFLNEGATEIFDAT 321
QY 301 QBPVPYAYKANWGLVDNKSFSVKAQWLKQNNFGGAMIWAILDDPTGSCFQCGKFPPLTS 360
DB 322 QBPVPYAYLGNWGVNDVNSFKLKAQWLKQNNLGGAVVWPLDMDDESGSFCHQGRPLTT 381
QY 361 TLNKAIGISTEGCTAP 376
DB 382 TLKRLNHNVSASCKP 397
RESULT 13
Q8VHG1 PRELIMINARY; PRT; 402 AA.
ID Q8VHG1
AC Q8VHG1;
DT 01-WAR-2002 (TrEMBLrel. 20, Created)
DT 01-WAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Toad pancreatic chitinase (EC 3.2.1.14).
GN YM2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=129/OLA;
RA Su W.B., Chang N.-C.A.;
"Genomic organization of the ym2 gene";
RT Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF461142; AAL66748.1; -;
DR InterPro; IPR001223; Glyco_hydro_18.

DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
SQ SEQUENCE 402 AA; 44975 MW; D935389319A59905 CRC64;
Query Match 55.1%; Score 1373; DB 11; Length 402;
Best Local Similarity 65.3%; Pred. No. 1.1e-90;
Matches 245; Conservative 61; Mismatches 69; Indels 0; Gaps 0;
QY 1 YNLICFTWTAQYRPGLSFKFDDINPCLCTHLIYAFAGMKNNEITTIWINDVTLKAPN 60
DB 22 YQLMCYITSWAKDRPTEGSGFKGNIDPCCLCTHLIYAFAGMKNNEITYLSEQDLRDYEALN 81
QY 61 DLKRNRSKLKTLAIGGWNFGTAPFTTWTSTONRQTFTSVIKFLQYGFGLDLDMWEY 120
DB 82 GLKDRNTELTLLAIGGWNFGTAPFTTWTSTONRQTFTSVIKFLQYGFGLDLDMWEY 141
QY 121 PGSRSGPPQDKHLFTVLKEMREAFQEAIESNRPRLMVTAAGVAGISNIQAGYIPELS 180
DB 142 PGSRSGPPKDKHLFSLVQEMRKAFEEESTLNHPRLLLTSTGAGFIDVIGSYKIPELS 201
QY 181 KYLDFIHWMTYDLHSGWEGYTGNSPLYKYPTETGSNAYLVNDVYVMYKNGNAPAEKLI 240
DB 202 QSLDYIQVMTYDLHDPKNGYTGNSPLYKSPYDICKSADLNVDSTIYWKDHGAASEKLI 261
QY 241 VGFPEYGHFTILRNPDSNGIGAPTSGDPAGAYTRQAGFWAYYEICTFLRSGATEVMDAS 300
DB 262 VGFPAVGHFTILSDPSKNGIGDPTVSAGPPGKYTNEQGLLAYFEICTFLNEGATEIFDAT 321
QY 301 QBPVPYAYKANWGLVDNKSFSVKAQWLKQNNFGGAMIWAILDDPTGSCFQCGKFPPLTS 360
DB 322 QBPVPYAYLGNWGVNDVNSFKLKAQWLKQNNLGGAVVWPLDMDDESGSFCHQGRPLTT 381
QY 361 TLNKAIGISTEGCTA 375
DB 382 TLKRLNHNVSASCKA 396
RESULT 14
Q90W34 PRELIMINARY; PRT; 488 AA.
ID Q90W34
AC Q90W34;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Toad pancreatic chitinase (EC 3.2.1.14).
GN TPCase.
OS Bufo japonicus (Japanese toad).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Bufo.
OX NCBI_TaxID=8387;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PANCREAS;
RA Oshima H.;
RL Thesis (2001), Department of Department of Biology, Waseda University,
Tokyo, Japan.
DR EMBL: AJ345054; CAC87888.1; -;
DR InterPro; IPR002086; Aldehyde_dehydr.
DR InterPro; IPR001579; Chitinase_18/2.
DR InterPro; IPR002557; Chitin bind_PerA.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF01607; CBM 14; 1.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
DR PROSITE; PS01095; CHITINASE_18; UNKNOWN_1.
KW Glycosidase; Hydrolase; Signal.
FT SIGNAL.
FT CHAIN 1 18 TOAD PANCREATIC CHITINASE.
SQ SEQUENCE 488 AA; 54319 MW; E75E1AD3CB4919 CRC64;
Query Match 55.0%; Score 1371; DB 13; Length 488;
Best Local Similarity 54.0%; Pred. No. 2e-90;

Matches 259; Conservative 55; Mismatches 128; Indels 38; Gaps 6;	
QY 1	YNLICVFTWAQVRPGSGFKDDINPCLCTHLIYAFAGQNNNEITTIENVDVTLKAFN 60
Db	19 YKLVCIYFTWNSQTRPDQKTVPGNIDPQCLTHLVAFATWNEHKAIPYENWDDVLYKQFN 78
QY 61	DLKNRNSKLTLLAIGGMNFGTAPFTTSTQNRQTFTITSVKFLRQYGPGLDLDWEY 120
Db	79 DLQKQKNLVTLAIGGMNFGTQKFTDMVASSGNRSIFIKSVIAYLRQNNFDGIDLDPEY 138
QY 121	PGSRGSPQDKHLFTVLVKEMREAFQEAIESNRPLMTAAVAGGISNIQAGYEIPELS 180
Db	139 PGSRGSPDEKORFTYLIQEMLDFAFNEEARSSGPLRLITAAVASAGKTIDAGYEITAKTG 198
QY 181	KYLDFTLHVTYDILHSGWEGYTGNSPLYKYPTETGSNAYLVNDVYMNWKNNGAPAEKLI 240
Db	199 QLLDFISVMTYDPHGGWDYQSGHNSPLCKGSTDYDGLQIFNHFAMNYWKNNGAPAEKLI 258
QY 241	VGPPEYGHFTFLRNPSTN--GIGAPTSGDGPAGAYTRQAGFWAYYEICTFLRSGATEVWD 298
Db	259 LGFPTYGRIF--RNPNPNCVGVIPVSGAGSAGPYTREAGFWAYYEICTWL--SGSTVKWI 315
QY 299	ASQEVPIYAKNEWLGYDNIKSFVKAQWLKQNNFGGAMIWALDLDFTGSCDQCKFPL 358
Db	316 PDQVPEYACKSNEWVGFNDQESYECKVRELKSGFGGAMVWALDLDFFGRFCNQGRYPL 375
QY 359	TSTLKAIGISTEGC-----TAPD--VPSEPTVTPPGSGSG 392
Db	376 INHLKSLLEGSTVNCPEICGGISIFTPTSATTTTTTAKPDCTTPEPPVTPPP----- 430
QY 393	GGSSGSSGGSGFCADKADGLYEPVADNFAWOCINGITYQHQAGLVFDTSNCNCNWP 452
Db	431 --VPPVIDVDPNFCVEKTDGLHVNPLNTKNFYICANGRTYSMKCADGLVFOASNCNCNWP 488

RESULT 15

Q8VH43

PRELIMINARY; PRT; 402 AA.

AC Q8VH43

DT 01-MAR-2002 (TrEMBLrel. 20, Created)

DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)

DE Putative secretory protein.

GN YM2.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BALB/C; TISSUE=STOMACH;

RA Yang S.C., Su W.B.;

RT "Tissue-specific Expression of Ym2 Protein.";

RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY065557; AAL57751.1;

DR InterPro; IPR001223; Glyco hydro 18.

DR Pfam; PF00704; Glyco_hydro_18; 1

DR ProDom; PD000471; Glyco_hydro_18; 1.

SQ SEQUENCE 402 AA; 44985 MW; D935382E89A5981A CRC64;

Query Match	54.9%;	Score 1368;	DB 11;	Length 402;
Best Local Similarity	65.1%;	Pred. No. 2.5e-90;		
Matches 244; Conservative 61; Mismatches 70; Indels 0; Gaps 0;				

QY 1	YNLICVFTWAQVRPGSGFKDDINPCLCTHLIYAFAGQNNNEITTIENVDVTLKAFN 60
Db	22 YQLMCTYTSNAKDRPTEGSGFKGNIDPCLCTHLIYAFAGMKNNEITLSEQDLRDYEALN 81
QY 61	DLKNRNSKLTLLAIGGMNFGTAPFTTSTQNRQTFTITSVKFLRQYGPGLDLDWEY 120
Db	82 GLKDRNTELTLLAIGKWKFGAPFFSWSTPQNRQTFIKSVIRFLRQYNFDFGLNLDWQY 141
QY 121	PGSRGSPQDKHLFTVLVKEMREAFQEAIESNRPLMTAAVAGGISNIQAGYEIPELS 180

Db	142 PCPRGSPKDKHLFSLVQENRKAFEEESTLNHPIRLLTSTGAGFIDVIKSGYKIPELS 201
QY 181	KYLDFTLHVTYDILHSGWEGYTGNSPLYKYPTETGSNAYLVNDVYMNWKNNGAPAEKLI 240
Db	202 QSLDIQVMTYDLDHPKNGYTGNSPLYKSPYDYGKSADLVNDSIITTYKDHGAASEKLI 261
QY 241	VGPPEYGHFTFLRNPSTNIGAPTSGDGPAGAYTRQAGFWAYYEICTFLRSGATEVWDAS 300
Db	262 VGFPAYGHTFILTSPSKNGIGDPTVSAGPPGKYTNEQGLLAYFEICTFLNEGATEIFDAT 321
QY 301	QEVPIYAKNEWLGYDNIKSFVKAQWLKQNNFGGAMIWALDLDFTGSCDQCKFPLTS 360
Db	322 QEVPIYALGNEWVGVDNVRSPFLKAQWLKQNNLGGAVVPLMDMDDFSGSFCHQGRPLTT 381
QY 361	TINKALGISTEGCTA 375
Db	382 TLKRLNVHSASCKA 396

Search completed: June 29, 2003, 21:01:46
Job time : 28.4083 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2003, 20:51:08 ; Search time 8.02367 seconds
(without alignments)
2336.500 Million cell updates/sec

Title: US-10-004-219b-9

Perfect score: 2493

Sequence: 1 YNLICYFTNWAQYRPLGSGF.....QQHCQAGLVFDTSCNCCNWP 452

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1051	42.2	721	1 OGP MOUSE	Q62010 mus musculus
2	1042.5	41.8	671	1 OGP MESAU	Q60557 mesocricetu
3	1013	40.6	678	1 OGP HUMAN	Q12889 homo sapien
4	1010	40.5	537	1 OGP BOVIN	Q28042 bos taurus
5	1009.5	40.5	623	1 OGP PAPAN	P36718 papio anubi
6	1003	40.2	539	1 OGP SHEEP	Q28542 ovvis aries
7	990.5	39.7	527	1 OGP PIG	Q28990 sus scrofa
8	987.5	39.6	383	1 C3L1 HUMAN	P36222 homo sapien
9	948.5	38.0	390	1 C3L2 HUMAN	Q15782 homo sapien
10	892.5	35.8	381	1 C3L1 MOUSE	Q61362 mus musculu
11	889.5	35.7	617	1 CHIT_CABEL	Q11174 caenorhabdi
12	863.5	34.6	554	1 CHIT_MANSE	P36362 manduca sex
13	810	32.5	504	1 CHIT BRUMA	P29030 brugia mala
14	484	19.4	699	1 CH11 BACCI	P20533 bacillus ci
15	466.5	18.7	423	1 CH11 APHAL	P32470 aphanocla
16	442.5	17.7	423	1 CH14 TRIHA	P48827 trichoderna
17	422	17.0	427	1 CH11_COCIM	P54196 coccidioid
18	365.5	14.7	499	1 CH1B_SERMA	P11797 serratia ma
19	351.5	14.1	563	1 CH1A_SERMA	P07254 serratia ma
20	341.5	13.7	1046	1 CHIT VIBFU	P06156 vibrio furn
21	328	13.2	550	1 CHIT NPVOF	Q10363 orgyia pseu
22	327.5	13.1	619	1 CHIT_STRLI	P36909 streptomyce
23	325.5	13.1	820	1 CH1A_ALTSO	P32823 alteromonas
24	315	12.6	610	1 CHIT_STRPL	P11220 streptomyce
25	300	12.0	551	1 CHIT NPVAC	P41684 autographa
26	233.5	9.4	1146	1 KTXA_KJULA	P09805 kluyveromyc
27	197.5	7.9	385	1 DIAC_HUMAN	Q01460 rattus norv
28	185.5	7.4	367	1 DIAC_RAT	Q01460 rattus norv
29	169.5	6.8	427	1 YAAH_BACSU	P37531 bacillus su
30	150	6.0	524	1 CH1D_BACCI	P27050 bacillus ci
31	136	5.5	491	1 YKCS_CABEL	P41996 caenorhabdi
32	131	5.3	597	1 CH1X_STROI	Q05638 streptomyce
33	130.5	5.2	812	1 PAED_ECOLI	P06970 escherichia

34	128	5.1	540	1	CH11_RHIOL	P29026 rhizopus ol
35	118.5	4.8	542	1	CH12_RHIOL	P29027 rhizopus ol
36	117.5	4.7	860	1	CH12_COCIM	P54197 coccidioid
37	114	4.6	36	1	C3L1_BOVIN	P30922 bos taurus
38	110.5	4.4	685	1	SC31_HUMAN	Q07837 homo sapien
39	106.5	4.3	346	1	RO21_KENIA	P51989 xenopus lae
40	104.5	4.2	958	1	AMYG_DEBOC	P22861 debaryomyce
41	103	4.1	666	1	MUR2_ENTHR	P39046 enterococcu
42	102.5	4.1	741	1	GUNS_CLOTM	P38686 clostridium
43	102	4.1	1385	1	CSAA_BACUD	Q45760 bacillus th
44	101.5	4.1	972	1	CTAI_BACCI	P94286 bacillus ci
45	101.5	4.1	2334	1	WAPA_BACSU	Q07833 bacillus su

ALIGNMENTS

RESULT 1
OGP_MOUSE STANDARD; PRT; 721 AA.
AC Q62010;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Oviduct-specific glycoprotein precursor (Oviductal glycoprotein)
DE (Oviductin) (Estrogen-dependent oviduct protein).
GN OVGPI OR OGP OR CHITS.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ICR; TISSUE=Oviduct;
RX MEDLINE=96115001; PubMed=7492880;
RA Sendai Y., Komiya H., Suzuki K., Onuma T., Kikuchi M., Hoshi H.,
RA Araki Y.;
RT "Molecular cloning and characterization of a mouse oviduct-specific
glycoprotein";
RL Biol. Reprod. 53:285-294(1995)
CC -!- FUNCTION: BINDS TO OOCYTE ZONA PELLUCIDA IN VIVO. MAY PLAY A ROLE
IN THE FERTILIZATION PROCESS AND/OR EARLY EMBRYONIC DEVELOPMENT.
CC -!- SUBCELLULAR LOCATION: Secretory granules.
CC -!- TISSUE SPECIFICITY: EPITHELIAL CELLS OF THE OVIDUCT.
CC -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.

THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL; D32137; BAA06863.1; -;
DR MGD; MGI:106661; Ovgp1.
DR InterPro; IPR001579; Chitinase_18/2.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR PROSITE; PS01095; CHITINASE_18; FALSE NEG.
KW Glycoprotein; Fertilization; Repeat; Signal.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 721
FT DOMAIN 486 632
FT OVIDUCT-SPECIFIC GLYCOPROTEIN.
FT 21 X 7 AA TANDEM REPEATS OF S-K-T-T-
FT [TAP]-G-(IV).
FT CARBOHYD 402 402 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 442 442 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 469 469 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 721 AA; 78807 MW; 37246C8F01665652 CRC64;
Query Match 42.2%; Score 1051; DB 1; Length 721;
Best Local Similarity 52.0%; Pred. No. 8.4e-68;

```
Matches 206; Conservative 61; Mismatches 119; Indels 10; Gaps 6;
QY 1 YNLICFTNWAQYRGLGSKPDDINPCLTHLIYAFAGQNNNEITTIENWD-VTLKAF 59
DB 22 YKLVCFYTNWAHSRPGFASIMPHDLDFLCTHLIFAFASNNQIVAKNLDENVLYPEF 81
QY 60 NDLKRNKSLKTLTLAGWNGFTAPFTTVMVSTSONQRTFTITSVIKELRGVDFGLDLWE 119
DB 82 NKLKERNRELKTLISLGGWNGFTSRFTMLSTLANKEKIDSVISFLRHGDFGLDLFL 141
QY 120 YPGSRGSPQDKHLFTVLVKEMREAFQEAIESRRRLMVTAAVAGGINSIQAGYEIPEL 179
DB 142 YPGLRGSPHPRNWFLELQFAFERALLQHPRLLSAAVSGIPSIHTSYDALL 201
QY 180 SKYLDHIVMTYDLHGSWEYTGNSPLKYPTETGSNAYLVNDVYNNYKNGAPAEKL 239
DB 202 GRRLDFINVLSDHGSWEKFTGHNSPLSLPDSKSSA-----YANNYWRKLGTPADKL 256
QY 240 IVGFPEYGHFTILRNPSDNGIGAPTSGDGPAGAYTRQAGFWAYEICTFLRSQATEVWDA 299
DB 257 IMGFTYGRNLYLLKESKGLQTASGPGSPGKYTKQAGFLAYEYVCSFQR-AKGHWID 315
QY 300 SQBPVYAYKANWGLYDNTKFSVKAQWLKQNNFGGAMIWDLDTGSCDQKGFPLT 359
DB 316 YQVVPYAFKGEWLGVDYDTSFSYKAMYKREHFGAMVWTLDMDDVRGTFCGNGPPLV 375
QY 360 STINKALGISTEGCTAPDVPSEPVTPPGSGSGGS 395
DB 376 HILNELL-VOTESNTP-LPQFWFTSSV-NASPGGS 408

RESULT 2
ID_OGP_MESAU STANDARD; PRT; 671 AA.
AC Q60527; Q60526;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Oviduct-specific glycoprotein precursor (Oviductal glycoprotein)
DE (Oviductin) (Estrogen-dependent oviduct protein) (ZP-O).
GN OVGPI OR OGP.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Oviduct;
RX MEDLINE=96115007; PubMed=7492686;
RA Suzuki K., Sendai Y., Osuma T., Hoshi H., Hiroi M., Araki Y.;
RT "Molecular characterization of a hamster oviduct-specific
RT glycoprotein.";
RL Biol. Reprod. 53:345-354 (1995).
RN [2]
RP SEQUENCE OF 22-671 FROM N.A.
RC TISSUE=Oviduct;
RX MEDLINE=96192955; PubMed=8607967;
RA Faquette Y., Merlen Y., Malette B., Bleau G.;
RT "Allelic polymorphism in the hamster oviductin gene is due to a
RT variable number of mucin-like tandem repeats.";
RL Mol. Reprod. Dev. 42:388-396 (1995).
RN [3]
RP SEQUENCE OF 14-671 FROM N.A., AND REVISIONS.
RC TISSUE=Oviduct;
RA Faquette Y.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 22-39.
RX MEDLINE=94058981; PubMed=8240241;
RA Malette B., Bleau G.;
RT "Biochemical characterization of hamster oviductin as a sulphated
RT zona pellucida-binding glycoprotein.";
```

```
RL Biochem. J. 295:437-445 (1993).
CC -!- FUNCTION: BINDS TO OOCYTE ZONA PELLUCIDA IN VIVO. MAY PLAY A ROLE
CC IN THE FERTILIZATION PROCESS AND/OR EARLY EMBRYONIC DEVELOPMENT.
CC MIGHT ACT AS A PROTECTIVE SECRETION INFLUENCING THE FIRST STEPS OF
CC THE REPRODUCTIVE PROCESS NECESSARY FOR THE NORMAL TRIGGERING OF
CC FERTILIZATION AND EARLY EMBRYONIC DEVELOPMENT.
CC -!- SUBCELLULAR LOCATION: Secretory granules.
CC -!- TISSUE SPECIFICITY: OVIDUCT.
CC -!- PTM: HIGHLY O-GLYCOSYLATED AND ALSO N-GLYCOSYLATED.
CC -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D32218; BAA06977.1; -.
CC EMBL; U15048; AAC53584.1; -.
CC InterPro; IPR001579; Chitinase_18/2.
CC InterPro; IPR001223; Glyco_hydro_18.
CC Pfam; PF00704; Glyco_hydro_18; 1.
CC ProDom; PD000471; Glyco_hydro_18; 1.
CC PROSITE; PS01095; CHITINASE_18; FALSE NEG.
KW Glycoprotein; Fertilization; Repeat; Signal.
FT SIGNAL 1 21
FT CHAIN 22 671 OVIDUCT-SPECIFIC GLYCOPROTEIN.
FT DOMAIN 490 609 8 X 15 AA TANDDEM REPEATS.
FT REPEAT 490 504 1.
FT REPEAT 505 519 2.
FT REPEAT 520 534 3.
FT REPEAT 535 549 4.
FT REPEAT 550 564 5.
FT REPEAT 565 579 6.
FT REPEAT 580 594 7.
FT REPEAT 595 609 8.
FT CARBOHYD 402 402 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 511 511 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 526 526 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 541 541 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 556 556 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 571 571 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 586 586 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 26 26 C -> A (IN REF. 4).
FT CONFLICT 33 33 H -> I (IN REF. 4).
FT CONFLICT 137 137 D -> G (IN REF. 2 AND 3).
FT CONFLICT 153 153 R -> Q (IN REF. 2 AND 3).
FT CONFLICT 165 165 F -> Y (IN REF. 2 AND 3).
FT CONFLICT 193 193 Q -> L (IN REF. 2 AND 3).
FT CONFLICT 531 545 MISSING (IN REF. 2 AND 3).
FT CONFLICT 595 595 T -> I (IN REF. 2 AND 3).
SQ SEQUENCE 671 AA; 73250 MW; BB57E0E514EC1972 CRC64;

Query Match 41.8%; Score 1042.5; DB 1; Length 671;
Best Local Similarity 51.4%; Pred. No. 3 1e-67;
Matches 204; Conservative 62; Mismatches 122; Indels 9; Gaps 5;
QY 1 YNLICFTNWAQYRGLGSKPDDINPCLTHLIYAFAGQNNNEITTIENWD-VTLKAF 59
DB 22 YKLVCFYTNWAHSRPGFASIMPHDLDFLCTHLIFAFASNNQIVAKNLDENVLYPEF 81
QY 60 NDLKRNKSLKTLTLAGWNGFTAPFTTVMVSTSONQRTFTITSVIKELRGVDFGLDLWE 119
DB 82 NKLKERNRELKTLISLGGWNGFTSRFTMLSTLANKEKIDSVISFLRHGDFGLDLFL 141
QY 120 YPGSRGSPQDKHLFTVLVKEMREAFQEAIESRRRLMVTAAVAGGINSIQAGYEIPEL 179
DB 142 YPGLRGSPHPRNWFLELQFAFERALLQHPRLLSAAVSGIPSIHTSYDALL 201
QY 180 SKYLDHIVMTYDLHGSWEYTGNSPLKYPTETGSNAYLVNDVYNNYKNGAPAEKL 239
```

Db 202 GRRLOFIVLSYDLHGSWEKSTGHSNPLSPEDPKSSA-----FAMNYWRNLGAPADKL 256
 QY 240 IVGFPEYGHFTILRNPSDNGIGAPTSGDGPAGAYTRQAGFWAYEICTFLRSGATEVWDA 299
 Db 257 LMGFPAYGRTFHLRESKNGLOASNGPASPCKYTKQAGFLAYEVCSPQR-AEKHWID 315
 QY 300 SEQVPYAYKANWGLVDNTKSFVKAQWLKQNNFGGAMIWALDDFTGSCDQGGKFPJT 359
 Db 316 HQVVPYAYKGEWGVYDDAVSFYKAMFVKEHFGGAMVWTLMDMDVRCGNGPFPPLV 375
 QY 360 STLNKALGISTEGCTAPDVPSPVTPPGSGGGSS 396
 Db 376 HILNELL-VRAEFNSTP-LPQFWFTLPVNSGPGSES 410

RESULT 3

OGP_HUMAN
 ID_OGP_HUMAN STANDARD; PRT; 678 AA.
 AC Q12889; Q15841;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Oviduct-specific glycoprotein precursor (Oviductal glycoprotein)
 DE (Oviductin) (estrogen-dependent oviduct protein).
 GN OVGPI OR OGP.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Oviduct;
 RA MEDLINE=95119256; PubMed=7819450;
 RA Arias E.B., Verhage H.G., Jaffe R.C.;
 RT "Complementary deoxyribonucleic acid cloning and molecular
 RT characterization of an estrogen-dependent human oviductal
 RT glycoprotein."
 RL Biol. Reprod. 51:685-694 (1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Jaffe R.C.;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: BINDS TO OOCYTE ZONA PELLUCIDA IN VIVO. MAY PLAY A ROLE
 CC IN THE FERTILIZATION PROCESS AND/OR EARLY EMBRYONIC DEVELOPMENT.
 CC -!- SUBCELLULAR LOCATION: Secretory granules.
 CC -!- TISSUE SPECIFICITY: OVIDUCT.
 CC -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.

This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

EMBL; U09550; AAB46946.1;
 EMBL; U58010; AAB04126.1;
 EMBL; U58001; AAB04126.1; JOINED.
 EMBL; U58002; AAB04126.1; JOINED.
 EMBL; U58003; AAB04126.1; JOINED.
 EMBL; U58004; AAB04126.1; JOINED.
 EMBL; U58005; AAB04126.1; JOINED.
 EMBL; U58006; AAB04126.1; JOINED.
 EMBL; U58007; AAB04126.1; JOINED.
 EMBL; U58008; AAB04126.1; JOINED.
 EMBL; U58009; AAB04126.1; JOINED.
 Genew; HGNC:8524; OVGPI.
 MIM; 603578;
 InterPro; IPR001579; Chitinase_18/2.
 DR InterPro; IPR001223; Glyco_hydro_18.
 DR Pfam; PF00704; Glyco_hydro_18; 1.
 DR ProDom; PD000471; Glyco_hydro_18; 1.

DR PROSITE; PS01095; CHITINASE 18; FALSE_NEG.
 KW Glycoprotein; Fertilization; Signal.
 FT SIGNAL 1 21
 FT CHAIN 22 678
 FT CARBOHYD 402 402
 FT CARBOHYD 441 441
 FT CARBOHYD 580 580
 FT CARBOHYD 596 596
 FT CARBOHYD 648 648
 FT CONFLICT 477 477
 FT CONFLICT 511 511
 FT CONFLICT 514 514
 FT CONFLICT 676 676
 SQ SEQUENCE 678 AA; 75421 MW; 245F2CEDCE92768B CRC64;
 Query Match 40.6%; Score 1013; DB 1; Length 678;
 Best Local Similarity 48.8%; Pred. No. 4,1e-65;
 Matches 203; Conservative 58; Mismatches 125; Indels 30; Gaps 5;
 QY 1 YNLICYFTNWAQYRPLGSGFKPDDINPCLCTHLIYAFAGMQNNEITTIEMND-VTLYKAF 59
 Db 22 HKLVCFYFTNWAHSRPGPASILPHDLDPFLCTHLIFAFASMNQIVAKDLQDEKILYPEF 81
 QY 60 NDLKWRNSKLTLLAIGWNFGTAPFTTMVSTSQNRQTFITSVIKFLQYGGDGLDWE 119
 Db 82 NKLKERNRELKTLISIGGNFCTSFRTTLMSTFANREKFIASVISLLRTHDFDGLDLPFL 141
 QY 120 YGSGSGSPQDKHLETVLVKERNREAFQEAIESNPRMLMVTAAVAGISNIQAGYIPEL 179
 Db 142 YFGLRGSPMDHRTWTFLELLFAFRKALLTMRPLLSAAVSGVPHIVQTSVDVRF 201
 QY 180 SKYLPFIHMTYDLHGSWEYTGNSPLYKYPTETGNSAYLNVDMYVWYKNGAPAEKL 239
 Db 202 GLLDFINVLSDHGSWERFTGHSPLFSLPEPKSSA-----YAMNYWRLGAPSEKL 256
 QY 240 IVGFPEYGHFTILRNPSDNGIGAPTSGDGPAGAYTRQAGFWAYEICTFLRSGATEVWDA 299
 Db 257 IMGITYGRTFRLKASKNGLOAQAIGPASPCKYTKQAGFLAYEVCSPQR-AEKHWID 315
 QY 300 SEQVPYAYKANWGLVDNTKSFVKAQWLKQNNFGGAMIWALDDFTGSCDQGGKFPJT 359
 Db 316 HQVVPYAYKGEWGVYDDAVSFYKAMFVKEHFGGAMVWTLMDMDVRCGNGPFPPLV 375
 QY 360 STLNKALGISTEGCTAPDVPSPVTPPGSGGGSS 392
 Db 376 HILNELL-VRAEFNSTP-LPQFWFTLPVNSGPGSES 431

RESULT 4

OGP_BOVIN
 ID_OGP_BOVIN STANDARD; PRT; 537 AA.
 AC Q28342;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Oviduct-specific glycoprotein precursor (Oviductal glycoprotein)
 DE (Oviductin) (estrogen-dependent oviduct protein) (Fragment).
 GN OVGPI OR OGP.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxId=9913;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 19-47.
 RC TISSUE=Oviduct;
 RX MEDLINE=94257768; PubMed=8199272;
 RA Sendai Y., Abe H., Kikuchi M., Satoh T., Hoshi H.;
 RT "Purification and molecular cloning of bovine oviduct-specific
 RT glycoprotein."
 RL Biol. Reprod. 50:927-934 (1994).
 CC -!- FUNCTION: BINDS TO OOCYTE ZONA PELLUCIDA IN VIVO. MAY PLAY A ROLE
 CC IN THE FERTILIZATION PROCESS AND/OR EARLY EMBRYONIC DEVELOPMENT.

modified and this statement is not removed. Usage by and for commercial/ entities requires a license agreement (see <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

EMBL: U43490; AAA85445.1; --
 DR InterPro; IPR001579; Chitinase 18/2.
 DR InterPro; IPR001223; Glyco_hydro_18.
 DR Pfam; PF00704; Glyco_hydro_18; 1.
 DR ProDom; PD000471; Glyco_hydro_18; 1.
 DR PROSITE; PS01095; CHITINASE 18; FALSE_NEG.
 KW Glycoprotein; fertilization; Signal.
 FT SIGNAL 1 21 BY SIMILARITY.
 FT CHAIN 22 527 OVIDUCT-SPECIFIC GLYCOPROTEIN.
 FT CARBOHYD 62 62 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 402 402 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 441 441 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 527 AA; 58519 MW; 31B78F49CA2363A2 CRC64;

Query Match 39.7%; Score 990.5; DB 1; Length 527;
 Best Local Similarity 48.2%; Pred. No. 1.2e-63;
 Matches 200; Conservative 66; Mismatches 124; Indels 25; Gaps 6;

QY	1	YNLCVPTNNAQVRPGLGSPKPDINCLCTHLIYAFAGMNNETIIEWNDVTL-VKAF	59
DB	22	HKLVCFVANNWAFSRGPASILPRDLDPFLCTHLVFAFAMNDSDQIVAKDARDSEIFYPEF	81
QY	60	NDLKRNRSKLKTLIAIGWNFGTAPFTTVMSTSONROTFTITSVIKFLQYCGFGLDLDDWE	119
DB	82	NQLKERNEKLTLLSIGWNFGTSRFTTMLSTFTNREKFTIRSAIGLRLTHGFDGLDLFFL	141
QY	120	YPSGRSGPPQDKHLFTVLVKEMREAFQEAIESNRPLMTAAVAGISINIQAGYEIPEL	179
DB	142	YPLGRGSPRDRWNFLPLEELLAFRRAQLTMRPRLLLSAAVSADPHVIQKAYDVRLL	201
QY	180	SKYLDFTHWMTYDLHGSEWGYTGENSEPLKYKYPETGSAVLYNDVYNNYKNGCAPAEKL	239
DB	202	GRLLDFINVLSDYDLHGSEWGYTGENSEPLKSLSDDPKSSA-----YTNWYRKLGAPPEKL	256
QY	240	IVGFPEYGHPTFILRNPSDNGIGAPTSGDGAGAYTROAGFWAYVEICTFLRSRGATEVMDA	299
DB	257	LMGFPTTGRTFRLLKASKNELGAENVGSPGKTKQAGFLAYEYVCSFVQR-AKKRWID	315
QY	300	SQEVPIYAYKANEWLGYDNIKFSVKAQWLKONNFGGAMIWAIDLDDFTGSGCQGGKPL-	358
DB	316	HQVVPYAYRGKMWGYDDISFSYKAFFIKKEHFGGAMVMTLDDVDRGTCGTRGPPFLV	375
QY	359	-----TSLNKALIGIST---EGCTAPD--VPSEPVITTPPGSGGGSSS	396
DB	376	YMLNDLLKAEVSSSTLSPGFGLSTVNSRSTCPESLAVTKDLTLDLILPLGGEA	430

RESULT 8

C3LI_HUMAN	STANDARD;	PRT; 383 AA.
ID	C3LI_HUMAN	
AC	P36222; P30923;	
DT	01-JUL-1993 (Rel. 26, Created)	
DT	01-JUN-1994 (Rel. 29, Last sequence update)	
DT	15-JUN-2002 (Rel. 41, Last annotation update)	
DE	Chitinase-3 like protein 1 precursor (Cartilage glycoprotein-39) (GP-	
DE	39) (39 kDa synovial protein) (YKL-40).	
GN	CHI3L1.	
OS	Homo sapiens (Human).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
OX	NCBI_TaxID=9606;	
RN	[1]	
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.	
RC	TISSUE=Cartilage;	
RX	MEDLINE=94064658; PubMed=8245017;	
RA	Hakala B.E., White C., Recklies A.D.;	
RT	"Human cartilage gp-39, a major secretory product of articular	
RT	chondrocytes and synovial cells, is a mammalian member of a chitinase	
RT	protein family."	

J. Biol. Chem. 268:25803-25810(1993).

[2]

RN SEQUENCE FROM N.A.

RX MEDLINE=97386591; PubMed=9244440;

RC TISSUE=Blood;

RA Rehli M., Krause S.W., Andressen R.;

RT "Molecular characterization of the gene for human cartilage gp-39

RT [CHI3L1], a member of the chitinase protein family and marker for

RT late stages of macrophage differentiation.";

RL Genomics 43:221-225(1997).

[3]

RN SEQUENCE OF 22-45.

RX MEDLINE=90328983; PubMed=2375755;

RA Nyirkos P., Golds E.E.;

RT "Human synovial cells secrete a 39 kDa protein similar to a bovine

RT mammary protein expressed during the non-lactating period.";

RL Biochem. J. 269:265-268(1990).

CC -!- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN THE CAPACITY OF CELLS TO

CC RESPOND TO AND COPE WITH CHANGES IN THEIR ENVIRONMENT.

CC -!- SUBCELLULAR LOCATION: Extracellular.

CC -!- TISSUE SPECIFICITY: PRESENT IN ARTICULAR CHONDROCYTES, SYNOVIAL

CC CELLS AS WELL AS IN LIVER. UNDETECTABLE IN MUSCLE TISSUES, LUNG,

CC PANCREAS, MONONUCLEAR CELLS, OR FIBROBLASTS.

CC -!- PTM: GLYCOSYLATED.

CC -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

CC or send an email to license@isb-sib.ch).

DR EMBL; M80927; AAA16074.1; -

DR EMBL; Y08374; CAA69661.1; -

DR EMBL; Y08375; CAA69661.1; JOINED.

DR EMBL; Y08376; CAA69661.1; JOINED.

DR EMBL; Y08377; CAA69661.1; JOINED.

DR EMBL; Y08378; CAA69661.1; JOINED.

DR PIR; S10677; S10677.

DR PIR; A33162; A33162.

DR Genew; HGNC:1932; CHI3L1.

DR MIM; 601525; -

DR InterPro; IPR001579; Chitinase_18/2.

DR InterPro; IPR001223; Glyco_hydro_18.

DR Pfam; PF00704; Glyco_hydro_18; 1_

DR Prodom; PD000471; Glyco_hydro_18; 1.

DR PROSITE; PS01095; CHITINASE_18; FALSE_NEG.

DR Glycoprotein; Signal.

FT SIGNAL 1 21

FT CHAIN 22 383 CHITINASE-3 LIKE PROTEIN 1.

FT CARBOHYD 60 60 N-LINKED (GLCNAC...) (POTENTIAL).

SQ SEQUENCE 383 AA; 42613 MW; 76ADD8298EECD21 CRC64;

```

RESULT 8
C3LI HUMAN
ID _C3LI HUMAN STANDARD; PRT; 383 AA.
AC P36222; P30923;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chitinase-3 like protein 1 precursor (Cartilage glycoprotein-39) (GP-39) (39 kDa synovial protein) (YKL-40).
GN CHI3L1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RP TISSUE=Cartilage;
RC MEDLINE=94064658; Pubmed=8245017;
RX Hakata B.E., White C., Recklies A.D.;
RT "Human cartilage gp-39, a major secretory product of articular chondrocytes and synovial cells, is a mammalian member of a chitinase protein family."
RT

```

[illegible]

Db 196 QHLDLISINTYDFHGAWRTTGHHSPLFRGQDASPDRENTDYAVGVNLRIGAPASKULV 255
QY 241 VGFPEYGHFTFILRNPDSNDGIGAPTSQDGPAGAYTROAGFWAYYEICTFLRSGATEVMDAS 300
Db 256 MGIPFGSFTLAS-SETGVGAPISGPIGPGFTKEAGTFLAYEICDFLR-CATVHRTIG 313
QY 301 QVVPYAYKANWGLVDNIKFSFKQAQLKQNNFGGAMIWALDIDFTGSCFQDG-KPPLT 359
Db 314 QQVPYATKGNQWGVDDQESVSKVQYLKDRQLAGAMVWALDIDFQGSFCQDLRFPLT 373
QY 360 STLNLKAL 366
Db 374 NAIKDAL 380

RESULT 9
C3L2 HUMAN
ID C3L2 HUMAN STANDARD; PRT: 390 AA.
AC Q15782; Q15783; Q15749;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Chitinase 3-like protein 2 precursor (YKL-39) (Chondrocyte protein 39).
GN Chi3L2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A. (SHORT AND LONG ISOFORMS).
RA Grossman A., Matsuyama T., Baker E., Waterhouse P., Sutherland G.R.,
RA Mak T.W.;
RT "Cloning of a novel lymphoid restricted human chitinase and
RT localization to lp13.3".
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
SEQUENCE FROM N.A. (SHORT ISOFORM).
RC TISSUE=Articular cartilage;
RX MEDLINE=96325055; PubMed=8702629;
RA Hu B., Trinh K., Figueira W.P., Price P.A.;
RT "Isolation and sequence of a novel human chondrocyte protein related
RT to mammalian members of the chitinase protein family.";
RL J. Biol. Chem. 271:19415-19420(1996).
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A SHORT FORM (SHOWN HERE) AND A
CC LONG FORM; MAY BE PRODUCED BY ALTERNATIVE SPLICING
CC -!- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN CHONDROCYTES, FOLLOWED
CC BY SYNIOCYTES, LUNG AND HEART. NOT DETECTED IN BRAIN, SPLEEN,
CC PANCREAS, AND LIVER.
CC -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U58515; AAB04534.1; -
CC EMBL; U58514; AAB04533.1; -
CC EMBL; U49835; AAC50597.1; ALT_INIT.
CC Genew; HGNC:1933; CHI3L2.
CC MIM; 601526; -
CC InterPro; IPR001579; Chitinase_18/2.
CC InterPro; IPR001223; Glyco_hydro_18.
CC Pfam; PF00704; Glyco_hydro_18; 1.
CC ProDom; PD000471; Glyco_hydro_18; 1.
CC PROSITE; PS01095; CHITINASE_18; FALSE NEG.
KW Hydrolase; Signal; Alternative splicing.
FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 390 CHITINASE 3-LIKE PROTEIN 2.
FT CARBOHYD 35 35 N-LINKED (GLCNAC. .) (POTENTIAL).

FT VARSPLIC 1 12 MGATTMOQKSLW -> SQQMLWSPDEVLVLCVCHTHIFT
FT AIWDSRTRFGSTDFSPHSI (IN LONG ISOFORM).
SQ SEQUENCE 390 AA; 43500 MW; 97B86A2F3AA35677 CRC64;
Query Match 38.0%; Score 948.5; DB 1; Length 390;
Best Local Similarity 47.7%; Pred. No. 8.2e-61;
Matches 176; Conservative 76; Mismatches 108; Indels 9; Gaps 5;
QY 1 YNLICYFTNWAQYRPGLSFKPDIDINPCLCTHLIYAFAGMQNNETITTEIWNDDVTLKAFN 60
Db 27 YKLCYFTNWSQDRQEPCKFTPENIDPFLCSHLIYFASIEENKVIKDKSEWMLYQTIN 86
QY 61 DLKRNNSKLTLAIGWNFGTAPTTWVSTSONQRTFITSVIKFLROYGDFGLDWEY 120
Db 87 SLKTKNPKLILLSIGYLFSGKGFHPMVDSTSRLEFINSIILFLRNHFDGLDVSMTY 146
QY 121 FSGRSPPQDKHLFTVLVKEMREAFQEALESNRPRLMVTAAVAGGISNIQAGYEIPELS 180
Db 147 PDQK----ENTH-FTVLHLEAEAFQDKFTKSTKERLLLTAGVSAGROMINDSYQVEKLA 201
QY 181 KYLDFIHVMYTDLHGSWEG--YTGENSPLYKYPTETGNSAYLVNVDYVMYKNGNGAPAEK 238
Db 202 KDLDFINLLSDFHGSWEKPLITGHNSPLSKGWQDRGPSSYNNVEYAVGYMTHKMPSEK 261
QY 239 LIVGPYGHFTILRNPDSNDGIGAPTSQDGPAGAYTROAGFWAYYEICTFLRSGATEVMD 298
Db 262 VVMGIPTYGHSTLAS-AETTVGAPASGPAAGPITSSGFLAYYEICQFLK-GAKITRL 319
QY 299 ASQVVPYAYKANWGLVDNIKFSFKQAQLKQNNFGGAMIWALDIDFTGSCDQKPEPL 358
Db 320 QQQVYPYAVKGNQWGVDDVKSMTETKQVFLKNLGLGAMISIDMDDFTGKSCNQGPYPL 379
QY 359 TSTLNKALG 367
Db 380 QVAVKRSLG 388

RESULT 10
C3L1 MOUSE
ID C3L1 MOUSE STANDARD; PRT: 381 AA.
AC Q61362;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Chitinase-3 like protein 1 precursor (Cartilage glycoprotein-39)
DE (GP-39) (BRP39 protein).
GN CHI3L1 OR BRP39.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Breast;
RX MEDLINE=95060797; PubMed=7970700;
RA Morrison B.W., Leder P.;
RT "neu and ras initiate murine mammary tumors that share genetic markers
RT generally absent in c-myc and int-2-initiated tumors.";
RL Oncogene 9:3417-3426(1994).
CC -!- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN THE CAPACITY OF CELLS TO
CC RESPOND TO AND COPE WITH CHANGES IN THEIR ENVIRONMENT (BY
CC SIMILARITY).
CC -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC

```
DR EMBL; X93035; CAAG3603.1; -.
DR MGD; MGI:1340899; Ch13L1.
DR InterPro; IPR001579; Chitinase_18/2.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
KW Glycoprotein; Signal.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 381 CHITINASE-3 LIKE PROTEIN 1.
FT CARBOHYD 60 60 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 381 AA; 43001 MW; EF658188184F0450 CRC64;

Query Match 35.0%; Score 892.5; DB 1; Length 381;
Best Local Similarity 46.0%; Pred. No. 8.1e-57;
Matches 169; Conservative 76; Mismatches 113; Indels 9; Gaps 5;

QY 1 YNLICVFTNAQYRPGSGFKPDDINPCLCTHLIYAFAGM-QNNEITTEWNDVTLKAF 59
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 22 YKLVCYFTSWSQYREGVGSFLPDIAIQFCLTHIYSFANISSDNMLSTWEWNDESNDKL 81
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 60 NDLKNSKLTLLAIGGNFGTAPFTTWSTSONRQTFITSVIKELROYGDFGLDWE 119
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 82 NKLKTRTNLTKLLSYGKWFGEKRESEASNTERTAFVRSVAPFLRSYGFGLDLWL 141
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 120 YPGSRGSPQDKHLFTVLVKEMREAFQEAIESNRPRMLVMTAAVAGGIGNIQAQYIPEL 179
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 142 YPLRL-----DKQYFTLKEINAEETKE-VQPGREKLLSALSAGKVAIDGYDIAOI 195
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 180 SKYLDPIHVMYDHLGWSGTYGENSPLYKYPTETGSNAYLVNDVYMYWKNAGAPAEKL 239
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 196 ACHLDPINLMTYDFHGVWRQITGHHSPLFGQKQDTRFDYSNVNVAVMIRLGAQAQSKL 255
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 240 IVGFPEYGHFTILRNPSDNGIGAPTSQDPAGATYQAGFWAYEICTELRSGATEVWDA 299
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 256 LMGIPTFGSKSFTLAS-SENQLGAPISGEGPLGRFTKEAGTFLAYEICTDLK-GAEVHRUS 313
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 300 SOBPVYAYKANEMWLDYNIKSFVKAQWLKQNNFGGAMIWAIDLDFTSQDQGFPLT 359
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 314 NEKVPATKGNQWGVYEHKESVKNVGLKELKLAGAMWALDLDLDFQTCQPKFFPLT 373
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 360 STLNKAL 366
DB : : ||
QY 374 NAIKDAL 380
DB : : ||

RESULT 11
CHIT_CABEL STANDARD; PRT; 617 AA.
AC Q11174;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Putative endochitinase (EC 3.2.1.14).
GN C04F6.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Nhan M.;
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
CC -|- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages of N-
CC acetyl-D-glucosamine polymers of chitin.
CC -|- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
CC HYDROLASES).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
```

```
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U42835; AAA83586.1; -.
DR WormPep; C04F6.3; CRO3923.
DR InterPro; IPR002557; Chitin bind PerA.
DR InterPro; IPR001579; Chitinase_18/2.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR Pfam; PF01607; CEM_14; 2.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR SMART; SM00494; ChtBD2; 2.
DR PROSITE; PS01095; CHITINASE_18; 1.
KW Hypothetical protein; Hydrolase; Glycosidase; Chitin degradation.
FT ACT SITE 179 179 PROTON DONOR (BY SIMILARITY).
FT DOMAIN 435 478 THR-RICH.
SQ SEQUENCE 617 AA; 66857 MW; DDAIDZAAAC0E54DA CRC64;

Query Match 35.7%; Score 889.5; DB 1; Length 617;
Best Local Similarity 39.6%; Pred. No. 2.5e-56;
Matches 190; Conservative 74; Mismatches 169; Indels 47; Gaps 14;

QY 5 CYFTNAQYRPGSGFKPDDINPCLCTHLIYAFAGM-QNNEITT-----IEWNVTL 55
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 57 CYFTNAQYRQGRKAFVPEYTPGLCTHLIFAF-GWNNADYTVRAYDPADLPNDWAGEGM 115
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 56 YKAFNDLKNRSKLTLLAIGGNFGTAPFTTWSTSONRQTFITSVIKELROYGDFGLD 115
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 116 YHRVNLKVTDTQLKTLISFGWSFGTALFQMAASSASRKVFIDSAITFTVWGFGLD 175
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 116 LDWEYFGSRGSPQDKHLFTVLVKEMREAFQEAIESNRPRMLVMTAAVAGGIGNIQAQY 175
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 176 IDWEYP-----SGATWANYVALVKELKAACEAGSTGKRLLVTAAGAAGPATIDAGYD 231
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 176 IPESKYLDPFIHVMYDHLGWSGTYGENSPLY-----KYPTETGSNAYLVNDVYMYWKN 232
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 232 IPNLAPNDFILLMSYDFGAWASLVGFNSPLYATTELPAAEWNG---WNVDSARYWQK 288
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 233 GAPASEKLIVGPEYGHFTILRNPSDNGIGAPTSQDPGA--GAYTRQAGFWAYEICTFLR 290
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 289 GMPKEKILVGMTYGRGWTLLNAS--AINPGTSG-SPAKITQYQVQAGVAYFECEMLA 345
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 291 SGATEVWDASQEVPIYAYKANEMWLDYNIKSFVKAQWLKQNNFGGAMIWAIDLDFTSQSF 350
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 346 NGATRYWDSQSQPVLYVQGNQWWSYDEESFANKWAYVKREGYGCAFVMTLDFDFDNAG- 404
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 351 CDQGK---FPLTSTLNKALG---ISTEG--CTAPVVPSEPVTT--PPGSGSGSGSGG-- 398
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 405 CSNSNGQLYPLISVIKELGVIIIPKGGVTTAPTATTATTATTGTPPMTSAVTTTTTAATT 464
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 399 -----SSGSGSGFCADKADGLYPVADDRNFWQCINGITYQQHCOAGLVDFDSCNCC 449
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 465 TTTRAATTTTASNTNVCSGKSDGFYPNSNNCGLFVLCUSSKSYSMSCPSGLQYSASLYKC 524
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

RESULT 12
CHIT_MANSE STANDARD; PRT; 554 AA.
AC P36362;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE Endochitinase precursor (EC 3.2.1.14).
OS Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata;
OC Ditrysia; Sphingioidea; Sphingidae; Sphinginae; Manduca.
OX NCBI_TaxID=7130;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93357793; PubMed=8353525;
RA Kramer K.J.; Corpuz L.; Choi H.K.; Muthukrishnan S.;
```



```
FT SIGNAL 1 22
FT CHAIN 23 504 ENDOCHITINASE.
FT DOMAIN 23 400 CATALYTIC.
FT DOMAIN 401 450 SER/THR-RICH (LINKER).
FT DOMAIN 407 448 3 X 14 AA APPROXIMATE TANDEM REPEATS.
FT ACT_SITE 148 148 PROTON DONOR (BY SIMILARITY).
SQ SEQUENCE 504 AA; 55971 MW; A78BE7BFBE3709B CRC64;

Query Match 32.5%; Score 810; DB 1; Length 504;
Best Local Similarity 35.2%; Pred. No. 9.5e-51;
Matches 168; Conservative 71; Mismatches 198; Indels 40; Gaps 6;

QY 5 CYFTNWAQYRGLSGFKPDINDPCLCTHLYAFAGMGN-NEITTIENNDVT-----LYK 57
DB 27 CYVTNWAQYRGLSGFKPLGNIPNGLCTHLYAFAGMDELGSKPFENWDETSKMGYS 86
QY 58 AFNDLKNRNSKLTLAIGAGNPGTAPFTMTWTSQNRQTFITSVKFLROYGFDGLDLD 117
DB 87 AVTKLRNTPGLKVLSSYGYNFGSAIFGTGIAKSAQKTERFTKSAIFLKRKNFDFGLD 146
QY 118 WEYVPGSGSPPODKHLFTVLVKEMREAFQEAIESNRPLMTVAAGGISNIQAGYEIP 177
DB 147 WEYVPG-----VAEEH--AKLVEAMKTAFAVEAKTSQKRLLLTAASACKGTIDGSYNE 200
QY 178 ELSKYLDIFHVMTYDLHGSWEGYTGNSPLYKYPTETGNSAVLNVDVVMYKNGNGAPAE 237
DB 201 SLGKNFDLLFLMSYDLHGSWEKNVDLHGLHPTKGEVSGIGIFNTEFAADYASWAKGMPKE 260
QY 238 KLVGPEYGHFTILRPNDSNGIGATSGDPAGATROAGFWAYEICTFLRSAGATEW 297
DB 261 KIIGIPMAQGWTLDPNPSETAIGAAASRPSSASKTNPAGGTASYWEICKYLKGGKQTV 320
QY 298 DASQEVYAYKANWELGYDNISFSVKAOWLKONNFGGAMIWAIDLDFTGSGCDGKFP 357
DB 321 HQEGVCAYVWGDQWGYNEETIRIMKWLKKEGKGAFIWAIDLDFTGSGCGKGP 380
QY 358 LSTLNKALGISTEGCTAPDVPSEPTVTPPGSGSGSGSGSGSGSGGFCADKA----- 410
DB 381 LLNAISSELEGESEN---PEITTEPSITETEAYETDEETSETAYDTEETSET 437
QY 411 -----DGLYPVADRNFAFWOCINGIITYQHCQAGLVDTSNCNCN 450
DB 438 ATTYDTEDEGECRPERDGLFPHPTDCHLFIOCANNIAYVMQCPATTFNFDAIKVCD 494

RESULT 14
CH11_BACCI STANDARD; PRT; 699 AA.
AC P20533;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE Chitinase A1 precursor (EC 3.2.1.14).
GN CH1A1.
OS Bacillus circulans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1397;
RN [1]_Strain=WL-12;
RP SEQUENCE FROM N.A.
RX STRAIN=WL-12;
RX MEDLINE=90368776; PubMed=2203782;
RA Watanabe T., Suzuki K., Oyanagi W., Ohnishi K., Tanaka H.;
RT "Gene cloning of chitinase A1 from Bacillus circulans WL-12 revealed
RT its evolutionary relationship to Serratia chitinase and to the type
RT III homology units of fibronectin.";
RL J. Biol. Chem. 265:15659-15665(1990).
RN [2]
RP MUTAGENESIS.
RX STRAIN=WL-12;
RX MEDLINE=93366760; PubMed=8103047;
RA Watanabe T., Kohori K., Miyashita K., Fujii T., Sakai H.,
RA Uchida M., Tanaka H.;
RT "Identification of glutamic acid 204 and aspartic acid 200 in
```

```
RT chitinase A1 of Bacillus circulans WL-12 as essential residues for
RT chitinase activity.";
RL J. Biol. Chem. 268:18567-18572(1993).
CC -!- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages of N-
CC acetyl-D-glucosamine polymers of chitin.
CC -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC -!- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
CC HYDROLASES).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M57601; AAA81528.1; -
CC PIR; A38368; A38368.
CC HSP; P02751; 1FNA.
CC InterPro; IPR003610; CBM 5.12.
CC InterPro; IPR001579; Chitinase_18/2.
CC InterPro; IPR003961; FN III.
CC InterPro; IPR003962; FNIII_repeat.
CC InterPro; IPR001223; Glyco_hydro_18.
CC Pfam; PF00041; fn3; 2.
CC Pfam; PF00704; Glyco_hydro_18; 1.
CC Pfam; PF02839; CBM 5.12; 1.
CC PRINTS; PR00014; FNTYPEIII.
CC ProDom; PD000471; Glyco_hydro_18; 2.
CC SMART; SM00495; ChtBD3; 1.
CC SMART; SM00060; FN3; 2.
CC PROSITE; PS01095; CHITINASE 18; 1.
CC Hydrolase; Glycosidase; Chitin degradation; Signal; Repeat.
FT SIGNAL 1 41
FT CHAIN 42 699 CHITINASE A1.
FT DOMAIN 42 460 CATALYTIC.
FT DOMAIN 465 549 FIBRONECTIN TYPE-III (R-1).
FT DOMAIN 560 644 FIBRONECTIN TYPE-III (R-2).
FT ACT_SITE 204 204 PROTON DONOR (PROBABLE).
FT MUTAGEN 200 200 D->N: DECREASE IN ACTIVITY.
FT MUTAGEN 200 200 D->E: NO CHANGE IN ACTIVITY.
FT MUTAGEN 204 204 E->D,Q: LOSS OF ACTIVITY.
SQ SEQUENCE 699 AA; 73677 MW; AC7C9B2E2987643 CRC64;

Query Match 19.4%; Score 484; DB 1; Length 699;
Best Local Similarity 27.4%; Pred. No. 3.2e-27;
Matches 130; Conservative 80; Mismatches 156; Indels 108; Gaps 18;

QY 1 YNLICYFTNWAQYRGLSGFKPDINDPCLCTHLYAFAGM----- 40
DB 44 YKIVGYPSWAAG--RNVNVAIDIDPTKVTHINYAFADICWNGHGNPDPSGPNPTWT 100
QY 41 -QNEITIE-----WNDVLYKAF-----NDLKNRNSKLTL 72
DB 101 CQNEKSQTINPVGNTIVLGDPWIDTG--KTFAGDTWQDFIAGNINQLNKLKQTNENLKI 158
QY 73 LAIGGNFGTAPFTMTWTSQNRQTFITSVKFLROYGFDGLDLDWEYVPGSG 127
DB 159 ISVGGWTWSNR-FSDVAATAATREVFANSADVFLKYNFDFGVLDLWEYVSGGLDGNKR 217
QY 128 PQDKHLFTVLVKEMREAFQEAIESNRPLMTVAAGGISNIQAGYIPELSKYLDPIH 187
DB 218 PEDKQNYTLTLLSKIREKLDAAAGVDGKKYLLTIA--SGASATYAANTELAIAIVDWIN 275
QY 188 VMTYDLHGSWEGYTGNSPLYKYPTET-----GSNAYLNVDVVMYKNG 233
DB 276 IMTYDFNGAWQKISAHNAPLNLYDPAASAAGVDPDANTFNVAAGAQHLLDA----- 325
QY 234 APAEKLIVGFPEYGHFTILRPNDSNGIGATSGDPAGATROAGFWAYEICTFLRS 291
DB 326 VPAKLVGLVFPYGRWGGCAAGNGQYQTCGTGSSVG--TWEAGSFDYDLEANYINKN 383
```


THIS PAGE BLANK (08/70)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 29, 2003, 21:10:28 ; Search time 1489.73 Seconds
(without alignments)
4913.891 Million cell updates/sec

Title: US-10-004-219b-9

Perfect score: 2493

Sequence: 1 YNLICYFTNWAQYRGLGSF.....QQHCQAGLVFTSCNCNWP 452

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=xlp
-Q/cgu2_1/USPTO_spool/US10004219/runat_24062003_160228_8173/app_query.fasta_1.1877
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -IOPCl=0 -IOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10004219 @CGN 1.1 4575 @runat_24062003_160228_8173 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:
1: em estba.*
2: em esthum.*
3: em estin.*
4: em estmu.*
5: em estov.*
6: em estpl.*
7: em estro.*
8: em htc.*
9: gb est1.*
10: gb est2.*
11: gb htc.*
12: gb est3.*
13: gb est4.*
14: gb est5.*
15: em estfun.*
16: em estom.*
17: gb gas.*
18: em gas hum.*
19: em gas inv.*
20: em gas pin.*
21: em gas vrt.*
22: em gas fun.*
23: em gas mam.*
24: em gas mus.*
25: em gas other.*
26: em gas pro.*
27: em gas rod.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2488	99.8	1533	11	AK008633	Mus muscu
2	2467	99.0	1535	11	AK008757	Mus muscu
3	2416	96.9	1534	11	AK008650	Mus muscu
4	2365.5	94.9	1530	11	AK007573	Mus muscu
5	1498	60.1	892	14	BQ231098	AGENCOURT
6	1452	58.2	900	12	BG867646	BG867646 602787435
7	1431	57.4	838	12	BG867815	BG867815 602786336
8	1416	56.8	809	12	BG866417	BG866417 602785352
9	1414	56.7	843	12	BG868207	BG868207 602784727
10	1400	56.2	872	14	BQ231064	AGENCOURT
11	1395	56.0	923	13	B1686954	B1686954 603313831
12	1394.5	55.9	830	12	BG865815	BG865815 602788220
13	1381.5	55.4	1007	9	AV072529	AV072529 AV072529
14	1363	54.7	857	12	BG871434	BG871434 602790636
15	1353	54.3	800	12	BG866853	BG866853 602786589
16	1349	54.1	763	12	BG872647	BG872647 602793705
17	1348	54.1	939	14	BQ219632	AGENCOURT
18	1346	54.0	764	12	BG869304	BG869304 602788819
19	1346	54.0	837	12	BG868741	BG868741 602787570
20	1341	53.8	894	12	BG86958	BG86958 602789826
21	1339	53.7	855	12	BG872136	BG872136 602792374
22	1335.5	53.6	949	12	BG868019	BG868019 602788381
23	1333	53.5	1100	12	BF302298	BF302298 602031269
24	1330.5	53.4	886	12	BG870840	BG870840 602791891
25	1330	53.3	808	12	BF534404	BF534404 602047909
26	1330	53.3	844	12	BF536293	BF536293 602051942
27	1324	53.1	757	12	BG867292	BG867292 602788635
28	1321.5	53.0	880	12	BG871468	BG871468 602790678
29	1321.5	53.0	1155	12	BG867473	BG867473 602788576
30	1315.5	52.8	836	12	BG869655	BG869655 602789716
31	1315.5	52.8	848	12	BG869143	BG869143 602789020
32	1313.5	52.7	874	12	BG865756	BG865756 602784345
33	1311	52.6	833	12	BG871337	BG871337 602790509
34	1309	52.5	830	12	BG869896	BG869896 602789435
35	1308.5	52.5	882	12	BG873488	BG873488 602791730
36	1306.5	52.4	788	12	BG873335	BG873335 602794304
37	1305	52.3	847	12	BG866775	BG866775 602786605
38	1305	52.3	856	12	BG870147	BG870147 602793390
39	1302	52.2	948	13	B1664452	B1664452 603290028
40	1299.5	52.1	845	12	BG865848	BG865848 602788264
41	1299	52.1	976	14	BQ225554	AGENCOURT
42	1297.5	52.0	863	12	BG872850	BG872850 602793973
43	1296	52.0	742	12	BG873143	BG873143 602794547
44	1293.5	51.9	878	12	BG868262	BG868262 602784787
45	1293	51.9	828	12	BG869811	BG869811 602789327

ALIGNMENTS

RESULT 1	AK008633	1533 bp	mRNA	linear	HTC 19-JAN-2002
LOCUS	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:2200003E03:chitinase, acidic, full insert sequence.				
DEFINITION	AK008633				
ACCESSION	AK008633.1	GI:12842941			
VERSION	HTC; CAP trapper.				
KEYWORDS	Mus musculus (strain:C57BL/6J) adult male stomach cDNA to mRNA, clone lib:RIKEN full-length enriched mouse cDNA library				
SOURCE	clone:2200003E03.				
ORGANISM	Mus musculus				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
REFERENCE	1				

Db 260 GACTTTGAAAAACAGGAACAGCAAACTGAAACCCCTCTGGCAATGGAGGCTTGAACCTT 319
QY 81 GlyThrAlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheIleThr 100
Db 320 GGAACCTGCTCTCTTCACTACCACTGGTTTCCACTTCTCAGAACCCGACACCTTCATTACC 379
QY 101 SerValIleLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyr 120
Db 380 TCAGTCATCAAAATTTCTGGCTCAGTATGGTTTGTATGGACTGGAGCTGGGAATAC 439
QY 121 ProGlySerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValLysGlu 140
Db 440 CCAGGCTCACGTGGAGGCCCTCTCAGGACAAAGCATCTTCACTGTCTCTGGTGAAGAA 499
QY 141 MetArgGluAlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThr 160
Db 500 ATGCGTGAAGCTTTTGAGCAGGAGGCTATTAGAGCAACAGGCCACGACTGATGGTTACT 559
QY 161 AlaAlaValAlaGlyIleSerAsnIleGlnAlaGlyTyrGluIleProGluLeuSer 180
Db 560 GCTGCTGTAGCTGTGGGATTTCCAACTCCAGGCTGGCTATGAGATCCTCGAATTTCT 619
QY 181 LysTyrLeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr 200
Db 620 AAGTACCTGGATTTCACTCATGTCTGATGATGATGATGATGATGATGATGATGATGAT 679
QY 201 ThrGlyGluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeu 220
Db 680 ACTGGGGAGAAAGTAGTCTCTTTACAAATACCTTACTGAGACTGGTAGCAATGCCATCTTC 739
QY 221 AsnValAspTyrValMetAsnTyrTrpLysAsnAsnGlyValAlaProAlaGluLysLeu 240
Db 740 AATGTGGATTATGTCATGAACTATTGGAAGACAACTGGAGCCCGAGCTGAGAAGCTCAT 799
QY 241 ValGlyPheProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIle 260
Db 800 GTTGGATTTCCAGAGTATGGACACACTTTCATCTCAGAAACCCCTCTCATTAATGGAAT 859
QY 261 GlyAlaProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrp 280
Db 860 GTTGCCCTTACTCTGTGTGTGATGCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 919
QY 281 AlaTyrTyrGluIleCysThrPheLeuArgSerGlyValAlaThrGluValTrpAspAlaSer 300
Db 920 GCCTACTATGAGATTGCACCTTCTGAGAGTGGAGCCACTGAGGTCTGGGATGCCTCC 979
QY 301 GlnGluValProTyrAlaTyrLysAlaAsnGluTrpLeuGlyTyrAspAsnIleLysSer 320
Db 980 CAAGAAGTGCCTATGCCTATAAGGCCAACGAGTGGCTTGGCTATGACAAATATCAAGAGC 1039
QY 321 PheSerValLysAlaGlnTrpLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTrpAla 340
Db 1040 TTCAGTGTTAAGGCTCAGTGGCTTAAGCAGAACAAATTTTGGAGGTGCCATGATCTGGGCC 1099
QY 341 IleAspLeuAspAspPheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSer 360
Db 1100 ATGACCTTGATGACTTCACTGGCTTCTTCTGTGATCAGGAAATTTCTCTGACTTCT 1159
QY 361 ThrLeuAsnLysAlaLeuGlyIleSerThrGluGlyCysThrAlaProAspValProSer 380
Db 1160 ACTTTTGAACAAAGCCCTTGGCATATCACTGAAGGTGTGCACAGCTCTCAGCTGCCTTC 1219
QY 381 GluProValThrThrProGlySerGlySerGlyGlyGlySerSerGlySerSer 400
Db 1220 GAGCCAGTGACTACTCTCCAGGAAGTGGAGTGGGGGTGGAGGTCCCGAGGAAGCTCT 1279
QY 401 GlyGlySerGlyPheCysAlaAspLysAlaAspGlyLeuTyrProValAlaAspAspArg 420
Db 1280 GGAGGCGATGGATTTCTGTGCCAAGCAGATGGCTCTTACCCTGTGCAGATGACAGA 1339
QY 421 AsnAlaPheTrpGlnCysIleAsnGlyIleThrTyrGlnGlnHisCysGlnAlaGlyLeu 440

Db 1340 AATGCTTTTGGCAGTCGATCAATGGAATCATACACCAGCAGCATTTGTCAAGAGGCTT 1399
QY 441 ValPheAspThrSerCysAsnCysCysAsnTrpPro 452
Db 1400 GTTTTGTATACAGCTGTAAATGCTGCAACTGGCCA 1435
RESULT 2
LOCUS AK008757
DEFINITION Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:2210020N23;chitinase, acidic, full insert sequence.
ACCESSION AK008757
VERSION AK008757.1 GI:12843148
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (strain:CS7BL/6J) adult male stomach cDNA to mRNA, clone:lib-RIKEN full-length enriched mouse cDNA library
clone:2210020N23.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komano, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Komano, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuda, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE 4
AUTHORS Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaudo, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Buit, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409 (6821), 685-690 (2001)
MEDLINE 21085660
PUBMED 11217851
REFERENCE 5 (bases 1 to 1535)
AUTHORS Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,

Qy	342	AspLeuAspAspPheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSerThr	361
Db	1103	GAACCTTGATGACATTCACCTGGCTCTTTCTGTGATCAGGAAAAATTCCTCTGACTTCTACT	1162
Qy	362	LeuAsnLysAlaLeuGlyIleSerThrGlnGlyCysThrAlaProAspValProSerGlu	381
Db	1163	TTGAACAAAGCCCTTGGCATATCCATGAAGTTGCACAGCTCTCTGACGTGCTTCGAG	1222
Qy	382	ProValThrThrProProGlySerGlySerGlyGlyGlySerSerGlyGlySerSerGly	401
Db	1223	CCAGTGACTACTCTCCAGCAAGTGGGGAGTGGGGTGAAGCTCCGAGGAGACTCTGGA	1282
Qy	402	GlySerGlyPheCysAlaAspLysAlaAspGlyLeuTyrProValAlaAspAspArgAsn	421
Db	1283	GGCAGTGGATTCTGTGCCGCAAGACGACAGATGGGCTCTACCCCTGTGGCAGATGACGAAAT	1342
Qy	422	AlaPheTyrGlnCysIleAsnGlyIleThrTyrGlnGlnHisCysGlnAlaGlyLeuVal	441
Db	1343	GCCTTTGGCAGTGTCATCATGTGATCATACATACAGCAGCATTTGTCAAGCAGGCTTGTT	1402
Qy	442	PheAspThrSerCysAsnCysCysAsnTyrPro 452	
Db	1403	TTTGATACCAAGCTGTAATTGCTGCAACTGGCCA 1435	
RESULT 3			
LOCUS	AK008650	1534 bp mRNA linear HTC 19-JAN-2002	
DEFINITION	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:2200008L17:chitinase, acidic, full insert sequence.		
ACCESSION	AK008650		
VERSION	AK008650.1 GI:12842971		
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (strain:C57BL/6J) adult male stomach cDNA to mRNA, clone lib:RIKEN full-length enriched mouse cDNA library		
	clone:2200008L17.		
ORGANISM	Mus musculus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus;		
REFERENCE	1		
AUTHORS	Carninci,P. and Hayashizaki,Y.		
TITLE	High-efficiency full-length cDNA cloning		
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)		
MEDLINE	99279253		
PUBMED	10349636		
REFERENCE	2		
AUTHORS	Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.		
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes		
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)		
MEDLINE	20499374		
PUBMED	11042159		
REFERENCE	3		
AUTHORS	Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T. and Hayashizaki,Y.		
TITLE	OKazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y. RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer		
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)		
MEDLINE	20530913		
PUBMED	11076861		
REFERENCE	4		
AUTHORS	Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y., Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S., Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I., Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R., Kadoya,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T., Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,		

Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Cardinale, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wyshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohatsu, S., and Hayashizaki, Y.

TITLE
Functional annotation of a full-length mouse cDNA collection
JOURNAL
Nature 409 (6821), 685-690 (2001)
MEDLINE
21085660
PUBMED
11217851
REFERENCE
5 (bases 1 to 1534)
AUTHORS
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Hori, F., Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Schriml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamanaka, I., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

TITLE
Direct Submission
JOURNAL
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-gsc@riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5', GAGAGAGAGAGATCCAGAGCTCTTTTTTTTTTTTTTTT 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5', GAGAGAGATCTCGATTAATTAATTAATTCCTCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI, 3' end: SstI.
Host: SOLR.

FEATURES
source
1..1534
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="PANTOM DB:2200008L17"
/db_xref="MGD:MGI:1894268"
/db_xref="taxon:10090"
/clone="2200008L17"
/sex="male"
/tissue_type="stomach"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
1..1534
misc_feature
/note="chitinase, acidic
data source:MGD, source key:MGI:1932052, evidence:ISS"
/db_xref="MGD:MGI:1932052"

BASE COUNT
378 a 383 c 374 g 399 t
ORIGIN

200 100 0

QY	321	PheSerVallysAlaGlnTrpLeuIysGlnAsnAsnPheGlyGlyAlaWetIleTTPAla	349	
Db	1038	TTCAAGTGTAAAGGCTCAGTGGCTTTAAAGCAGAAACAATTTTGGAGGTGCCATGATCTGGGGC	109	
QY	341	IleAspLeuAspAspPheThrGlySerPheCysAspGlnGlyIysPheProLeuThrSer	360	
Db	1098	ATTGACCTTGATGACTTTCACCTGGCTCTTTCTGTGATCAGGGAAAATTTCTCTGACTTCT	115	
QY	361	ThrLeuAsnIysAlaLeuGlyIleSerThrGlnGlyCysThrAlaProAspValProSer	380	
Db	1158	ACTTTGGAACAAGCCCTTGGCATATCCACTGAAGGTTGCACAGCTCTCTGACGTGCCTTC	121	
QY	381	GluProValThrThrProGlySerGlySerGlyGlyGlySerSerGlyGlySerSer	400	
Db	1218	GAGCCAGTGACTACTCTCCAGGAAGTGGGAGTGGGGTGGGAAGCTCCCGAGAGACTCT	127	
QY	401	GlyGlySerGlyPheCysAlaAspIysAlaAspGlyLeuTrpValAlaAspAspArg	420	
Db	1278	GGAGSCAGTGGATCTGTGCCGCAACAGCAGATGGCTCTACCTCTGGCAGATGACAGA	133	
QY	421	AsnAlaPheTrpGlnCysIleAsnGlyIleThrTyrGlnGlnHisCysGlnAlaGlyLeu	440	
Db	1338	AATGCTTTTGGCAGTGCATCAATGGAATCACATACCAGCAGCAITTTGTCGAAGCAGGCTT	139	
QY	441	ValPheAspThrSerCysAsnCysCysAsnTrpPro	452	
Db	1398	GTTTTGTATACAGCTGTAATGTGCACTGGCCA	1433	
RESULT 4	AK007573	1530 bp	linear	HTC 19-JAN-2000
LOCUS	Mus musculus	10 day old male pancreas cDNA, RIKEN full-length		
DEFINITION	enriched library, clone:1810021D23:chitinase, acidic, full insert			
ACCESSION	AK007573			
VERSION	AK007573.1	GI:12841201		
KEYWORDS	HTC; CAP trapper.			
SOURCE	Mus musculus (strain:C57BL/6J) 10 day old male pancreas cDNA to			
	mRNA, clone_lib:RIKEN full-length enriched mouse cDNA library			
	clone:1810021D23.			
ORGANISM	Mus musculus			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Sciurognathi; Muridae; Murinae; Mus.			
TITLE	Carninci, P. and Hayashizaki, Y.			
JOURNAL	Normalization and subtraction of cup-trapper-selected cDNAs to			
MEDLINE	Meth. Enzymol. 303, 19-44 (1999)			
PUBMED	99279253			
REFERENCE	10349636			
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,			
TITLE	Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.			
JOURNAL	Normalization and subtraction of cup-trapper-selected cDNAs to			
MEDLINE	prepare full-length cDNA libraries for rapid discovery of new genes			
PUBMED	Genome Res. 10 (10), 1617-1630 (2000)			
REFERENCE	20499374			
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,			
TITLE	Konno, H., Akiyama, J., Nishi, K., Kiteunai, T., Tashiro, H., Itoh, M.,			
JOURNAL	Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,			
MEDLINE	Yanamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,			
PUBMED	Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, S., Watahiki, M.,			
REFERENCE	Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,			
AUTHORS	Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.			
TITLE	RIKEN integrated sequence analysis (RISA) system-384-format			
JOURNAL	sequencing pipeline with 384 multiplexed capillary sequencer			
MEDLINE	Genome Res. 10 (11), 1757-1771 (2000)			
PUBMED	20530913			
REFERENCE	11076861			
AUTHORS	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,			


```

915 GCCTACTATGAGATTGACCTTTCTGAGAAGTGGAGCCACTGAGTCTGGGATGCTCC 974
301 GlnGluValProTyrAlaTyrLeuAlaAsnGluTyrLeuGlyTyrAspAsnIleTyrSer 320
975 CAAGAAGTGGCCCTATGCTATAGGCCAAGCAGTGGCTTGGCTATGACATATCAAGAGC 1034
321 PheSerValTyrAlaGlnTyrLeuAlaAsnGlnAsnPheGlyGlyAlaMetIleTyrAla 340
1035 TTCAGTGTAAAGGCTCAGTGGCTTAAGCAGAACAAATTTGGAGGTGCCATGCTGGGCC 1094
341 IleAspLeuAspAspPheThrGlySerPheCysAspGlnGlyLysPheProLeuThrSer 360
1095 ATTGACCTTGATCACTCTCTTCTGTGATCAGGAAATTTCTCTGACTTCT 1154
361 ThrLeuAsnIleAlaLeuGlyIleSerThrGluGlyCysThrAlaProAspValProSer 380
1155 ACTTTGAACAAAGCCCTTGGCATATCCACTGAAGTTGCACAGCTCTGACGGCTTCC 1214
381 GluProValThrThrProProGlySerGlySerGlyGlyGlySerSerGlyGlySerSer 400
1215 GAGCCAGTGACTACTCTCCAGAAAGTGGAGTGGGGTGGAGCTCCGGAGGAAGCTCT 1274
401 GlyGlySerGlyPheCysAlaAspLysAlaAspGlyLeuTyrProValAlaAspAspArg 420
1275 GGAGGCAGTGGATTCTGTGCCGACAAAGCAGATGGCTCTTACCTGTGGCAGATGACAGA 1334
421 AsnAlaPheThrGlnCysIleAsnGlyIleThrTyrGlnGlnHisCysGlnAlaGlyLeu 440
1335 AATGCTTTTGGCAGTGATCAATGAATGATACATACAGAGCATTTGTCAAGCAGGGCTT 1394
441 ValPheAspThrSerCysAsnCysCysAsnTyrPro 452
1395 GTTTTGTATACCACTGTAATTTCTGCACTGGCCA 1430

RESULT 5
BQ231098
LOCUS BQ231098
DEFINITION AGENCOURT 7578138 NCI CGAP St1 Mus musculus cDNA clone
IMAG:6051580 5', mRNA sequence.
ACCESSION BQ231098
VERSION BQ231098.1 GI:20412498
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 892)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LLAM13305
High quality sequence stop: 772.
Location/Qualifiers
1..892
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:6051580"
/clone_lib="NCI CGAP St1"
/lab_host="DH10B (Tr-resistant)"
/notes="Organ: stomach; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.77 kb. Library constructed by Life
Technologies. Note: this is a NCI CGAP Library."
FEATURES
Source

```

```

BASE COUNT 219 a 226 c 216 g 230 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 1.61e-148 Length: 892
Score: 1498.00 Matches: 279
Percent Similarity: 98.94% Conservative: 0
Best Local Similarity: 98.94% Mismatches: 1
Query Match: 60.09% Indels: 2
DB: 14 Gaps: 0
US-10-004-219B-9 (1-452) x BQ231098 (1-892)
QY 1 TyrAsnLeuIleCysTyrPheThrAsnTyrAlaGlnTyrArgProGlyLeuGlySerPhe 20
Db 36 TACATCTGATATGCTATTTTACCACTGGGCCAGTATCGGCCAGGCTCGGGAGCTTC 95
QY 21 LysProAspAspIleAsnProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet 40
Db 96 AAGCCTGATGACATTAACCCCTGCTGTGATCTACCTGATCTATGCCTTTTCTGGGATG 155
QY 41 GlnAsnAsnGluIleThrThrIleGluTyrPheAspValThrLeuTyrIleValaPheAsn 60
Db 156 CAGAACATGATGATCACCACCATAGATGGAATGATGTTACTCTCTATAAAGCTTTCAAT 215
QY 61 AspLeuLysAsnArgAsnSerLysLeuLysThrLeuLeuAlaIleGlyGlyTyrAsnPhe 80
Db 216 GACTTGAAAACAGAGAACAGCACAACCTTGGCAATTTGGAGCTGGAACTTT 275
QY 81 GlyThrAlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheIleThr 100
Db 276 GGAACCTGCTCTTTTCACTACCATGCTTTTCCATCTCTCAGAACCCGACAGCTTCATTAC 335
QY 101 SerValIleLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTyrGluTyr 120
Db 336 TCAGTCATCAAAATTTCTGCGTCAGTATGGGTTTGAATGAGCTGGACCTGGAGTGGAAATAC 395
QY 121 ProGlySerArgGlySerProGlnAspLysHisLeuPheThrValLeuValLysGlu 140
Db 396 CCAGGCTCAGCTGGGGAGCCCTCTCAGGACCAAGCATCTTCTCAGCTGCTCTGGTGAAGAA 455
QY 141 MetArgGluAlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThr 160
Db 456 ATGCGTGAAGCTTTTGAAGCAGGAGCTATTGAGACCAACAGGCCAGACTGATGCTTACT 515
QY 161 AlaAlaValAlaGlyGlyIleSerAsnIleGlnAlaGlyTyrGluIleProGluLeuSer 180
Db 516 GCTGCTGTAGCTGGTGGGATTTCCAAACATCCAGGCTGGCTATGAGATCCCTGAACTTTCT 575
QY 181 LysTyrLeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTyrProGluGlyTyr 200
Db 576 AAGTACCTGGATTTTCATCTCATGTCATGACATATGACCTCCATGGCTCTGGAGGGCTTAC 635
QY 201 ThrGlyGluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeu 220
Db 636 ACTGGGAGAGATAGTCTCTTTTACAAATACCTACTAGAGCTGGTAGCAATGCCCTACCTC 695
QY 221 AsnValAspTyrValMetAsnTyrTyrPheAsnAsnGlyAlaProAlaGluLysLeuIle 240
Db 696 AATGTGGATTTATGTCATGAATCTTTGGAGAACCAATGAGGCCAGCTGAGAGGCTCAAT 755
QY 241 ValGlyPheProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIle 260
Db 756 GTTGGATTTCCAGAGTATGGACACACTTTCATCTCGAAGAACCCCTCTGTATATGAAT 815
QY 261 GlyAlaProThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGln-AlaGly-PheThr 280
Db 816 GGTGCCCCCTACTCTGTGTGATGGCCCTCTGCTGGCCCTATACAGACAGGGCTGGGTCT 875
QY 280 tp 280
Db 876 GG 877

```


/lab_host="DH10B (T1 phage-resistant)"
/notes="Organ: salivary gland; Vector: pCMV-SPORT6; Site: 1:
NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo
dT. Average insert size 1.3 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT: 209 a 211 c 198 g 220 t

ORIGIN

Alignment Scores:
Pred. No.: 1 1.94e-141 Length: 838
Score: 1431.00 Matches: 263
Percent Similarity: 98.88% Conservative: 2
Best Local Similarity: 98.13% Mismatches: 3
Query Match: 57.40% Indels: 0
DB: 12 Gaps: 0

US-10-004-219B-9 (1-452) x BG867815 (1-838)

```
QY 1 TyrAsnLeuileCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPhe 20
DB 34 TACAATCTGATATGCTATTTCACCAACTGGGCCAGTATCGGCCAGGTCTGGGGAGCTTC 93
QY 21 LysProAspAspIleAsnProCysLeuCysThrHisLeuileTyrAlaPheAlaGlyMet 40
DB 94 AAGCCTGATGACATTAACCCCTCCCTGTGTACTCAGCTGATCTATGCCCTTTGCTGGGATG 153
QY 41 GlnAsnAsnGluileThrThrileGluTrpAsnAspValThrLeuTyrLysAlaPheAsn 60
DB 154 CAGAACATGAGATACCCACCATAGATGGAATGATGTTACTCTCTATAAAGCTTTCAT 213
QY 61 AspLeuLysAsnArgAsnSerLysLeuLysThrLeuLeuAlaileGlyGlyTrpAsnPhe 80
DB 214 GACTTGAAAAACAGGAACAGCAAACTGAAACCCCTCTGGCAATTTGGAGCTGGAACCTT 273
QY 81 GlyThrAlaProPheThrMetValSerThrSerGlnAsnArgGlnThrPheileThr 100
DB 274 GGAACCTGCTCTTTTCACTACCATGGTTTCCACTCTCAGAACCCGCGAGACCTTCATTACC 333
QY 101 SerValileLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyr 120
DB 334 TCAGTCATCAAAATTTCTGGCTCAGTATGGTTGATGGACTGGACCTGGACTGGGATAC 393
QY 121 ProGlySerArgGlySerProGlnAspLysHisLeuPheThrValLeuValLysGlu 140
DB 394 CCAGGCTCAGCTGGGAGCCCTCCTCAGSACAAGCATCTCTCACTGCTCTGGTGAAGGAA 453
QY 141 MetArgGluAlaPheGluGlnGluAlaileGluSerAsnArgProArgLeuMetValThr 160
DB 454 ATCGGTGAAGCTTTTGAGCAGGAGGCTATTGAGAGCAACAGGCCCAAGACTGATGGTTACT 513
QY 161 AlaAlaValAlaGlyGlyLysSerAsnileGlnAlaGlyTyrGluileProGluLeuSer 180
DB 514 GCTGCTGTAGCTGGTGGATTTCCACATCCAGGCTGGCTATGAGATCCTTGACCTTCT 573
QY 181 LysTyrLeuAspPheileHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr 200
DB 574 AAGTACCTGGATTTTCATGATGTCATGACATATGACCTCATGGCTCTCTGGAGGGCTAC 633
QY 201 ThrGlyGluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeu 220
DB 634 ACTGGGGAGAATAGTCTCTTTTACAATACCTTACTGAGACTGGTAGCAATGCTTACCTC 693
QY 221 AsnValAspTyrValMetAsnTyrTrpLysAsnAsnGlyValaProAlaGluLysLeuile 240
DB 694 AATGTGATATATGTCATGAATATTTGGAAGAACATGAGGCCCACTGAGAGCTCAT 753
QY 241 ValGlyPheProGluTyrGlyHisThrPheileLeuArgAsnProSerAspAsnGlyile 260
DB 754 GTTGGATTCACAGATTGGACACACTTTCATCCTCAGAAACCCCTCTGATATGATTT 813
QY 261 GlyAlaProThrSerGlyAspGly 268
DB 814 GGTGCCCTTACTCTGTGTGAATGTC 837
```

RESULT 8

BG866417

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BG866417 809 bp mRNA linear EST 29-MAY-2001
602785352F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4911421 5',
mRNA sequence.

BG866417 GI:14216957

BG866417

EST.

house mouse.

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 809)

NIH-MGC <http://mgs.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgpbs-remail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM10813 row: h column: 14

High quality sequence stop: 804.

Location/Qualifiers

FEATURES

source

1..809

/organism="Mus musculus"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone="IMAGE:4911421"

/clone_lib="NCI CGAP_SG2"

/lab_host="DH10B (T1 phage-resistant)"

/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site: 1:

NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo

dT. Average insert size 1.3 kb. Constructed by Life

Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT 193 a 200 c 208 g 207 t 1 others

ORIGIN

Alignment Scores:

Pred. No.: 7,21e-140 Length: 809
Score: 1416.00 Matches: 264
Percent Similarity: 98.14% Conservative: 0
Best Local Similarity: 98.14% Mismatches: 4
Query Match: 56.80% Indels: 1
DB: 12 Gaps: 0

US-10-004-219B-9 (1-452) x BG866417 (1-809)

```
QY 71 ThrLeuLeuAlaileGlyGlyTrpAsnPheGlyThrAlaProPheThrThrMetValSer 90
DB 2 ACCCTCTGGCAATTTGGAGCTGGAACTTTGGAACCTCTTCACTACCACTGGTTCC 61
QY 91 ThrSerGlnAsnArgGlnThrPheileThrSerValileLysPheLeuArgGlnTyrGly 110
DB 62 ACTTCTCAGAACCGCCAGACCTTCATTACCTCAGTCATCAAAATTTCTGCTCAGTATGG 121
QY 111 PheAspGlyLeuAspLeuAspTrpGluTyrProGlySerArgGlySerProGlnAsp 130
DB 122 TTTGATGACTGGACTGGACTGGGAATACCCAGGCTCAGCTGGGAGCCCTCTCTCAGGAC 181
QY 131 LysHisLeuPheThrValLeuValLysGluMetArgGluAlaPheGluGlnGluAlaile 150
DB 182 AAGCATCTCTTCACTGCTCTGCTGAAGGAAATGCGTGAAGCTTTTTCAGCAGGAGCTATT 241
QY 151 GluSerAsnArgProArgLeuMetValThrAlaAlaValAlaGlyGlyileSerAsnile 170
DB 242 GAGAGCAACAGGCCCAAGACTGATGTTACTGCTGTCTAGCTGGTGGGATTTTCCACATC 301
```

171 GlnAlaGlyThrGluLeuProGluLeuSerTyrLeuAspPheHisValMetThr 190
 Db CAGGCTGGCTATGAGATCCCTGAACCTTTCTAAGTACCTGGATTTCATCATGTCATGACA 361
 QY TyrAspLeuHisGlySerTyrGluGlyThrGluGluAsnSerProLeuTyrIlyThr 210
 Db TATGACCTCCATGGCTCTCTGGAGGGCTACACTGGGGAGAAATAGTCCTCTTTACAAATAC 421
 QY ProThrGluThrGlySerAsnAlaTyrLeuAsnValAspTyrValMetAsnTyrIlyThr 230
 Db CCTACTGAGACTGGTAGCAATGCTCACTCAATGTGGATTATGTCATGACTATTTGGAG 481
 QY AsnAsnGlyAlaProAlaGluLysLeuLeuValGlyPheProGluTyrGlyHisThrPhe 250
 Db AACATGGAGGCCAGCTGAGAGGCTCATTTGTGGATTCCAGAGTATGGACACACCTTC 541
 QY IleLeuArgAsnProSerAspAsnGlyIleGlyAlaProThrSerGlyAspGlyProAla 270
 Db ATCCCTGAGAAACCCCTCTGATTAATGGAATGGTGGCTTACCTCTGGTGATGGCCCTGCT 601
 QY GlyAlaTyrThrArgGlnAlaGlyPheThrPalatTyrGluIleCysThrPheLeuArg 290
 Db GGGCCCTATACAGACAGCTGGTCTGGGCTACTATGATGATGACCTTTCTGAGA 661
 QY SerGlyAlaThrGluValTrp-AspAlaSerGlnGluValProTyrAlaTyrLysAlaAs 310
 Db AGTGAGGACCTAGAGTCTGGGATGCTCCCAAGAGTGGCCCTATGCTATAGAGCCAA 721
 QY nGluTrpLeuGlyTyrAspAsnIleGlySerPheSerValLysAlaGlnTrpLeuLysG 330
 Db CGAGTGGCTTGGCTATGACATTCATCAAGAGCTTCAGTGTAAAGGTCAGTGGCTTAAACA 781
 QY nAsnAsnPheGlyGlyAlaMetIle 338
 Db GAACAATTTGGAGTGCCATGATC 806
 RESULT 9
 LOCUS BG868207
 DEFINITION 602784727F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4910958 5',
 mRNA sequence.
 ACCESSION BG868207
 VERSION BG868207.1 GI:14218747
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: c9apbs-remail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM10812 row: e column: 07
 High quality sequence stop: 792.
 Location/Qualifiers
 1..843
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:4910958"
 /clone_lib="NCI_CGAP_SG2"
 /lab_host="DH10B (T1 phage-resistant)"
 /note="Organ: salivary gland; Vector: pCMV-SPORT6; Site 1:
 NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo

dt. Average insert size 1.3 kb. Constructed by Life
 Technologies. Note: this is a NCI_CGAP Library."
 BASE COUNT 204 a 207 c 217 g 215 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.26e-139 Length: 843
 Score: 1414.00 Matches: 275
 Percent Similarity: 96.83% Conservative: 0
 Best Local Similarity: 96.83% Mismatches: 4
 Query Match: 56.72% Indels: 6
 Gaps: 0
 DB: 12
 US-10-004-219B-9 (1-452) x BG868207 (1-843)
 QY 111 PheAspGlyLeuAspLeuAspTyrGluTyrProGlySerArgGlySerProProGlnAsp 130
 Db 2 TTTGATGGACTGGACCTGGACTGGGAATACACAGGCTCACGTGGAGGCCCTCTCAGGAC 61
 QY 131 LysHisLeuPheThrValLeuValLysGluMetArgGluAlaPheGluGlnGluAlaIle 150
 Db 62 AAGCATCTCTTCACTGTCTGCTGGTGAAGAAATGCGTGAAGCTTTTGACAGAGGCTATT 121
 QY 151 GluSerAsnArgProArgLeuMetValThrAlaValAlaGlyGlyIleSerAsnIle 170
 Db 122 CAGAGCAACAGGCCACAGCTGATGCTGCTGCTAGCTGGTGGGATTTCCAAATC 181
 QY 171 GlnAlaGlyTyrGluIleProGluLeuSerIlyThrLeuAspPheIleHisValMetThr 190
 Db 182 CAGGCTGGCTATGAGATCCCTGAACTTTCTAAGTACCTGGATTTTCATCATGTGTAAGA 241
 QY 191 TyrAspLeuHisGlySerTyrGluGlyTyrThrGlyGluAsnSerProLeuTyrIlyThr 210
 Db 242 TATGACCTCCATGGCTCTCTGGAGGGCTACACTGGGAGAGATAGTCTCTTTTACAAATC 301
 QY 211 ProThrGluThrGlySerAsnAlaTyrLeuAsnValAspTyrValMetAsnTyrTrpLys 230
 Db 302 CCTACTGAGACTGGTAGCAATGCTCACTCAATGGGATTTATGTCATGAACCTATTGGAG 361
 QY 231 AsnAsnGlyAlaProAlaGluLysLeuIleValGlyPheProGluTyrGlyHisThrPhe 250
 Db 362 AACATGGAGGCCAGCTGAGAGGCTCAATGTTGGATTCCAGAGTATGGACACACCTTC 421
 QY 251 IleLeuArgAsnProSerAspAsnGlyIleGlyAlaProThrSerGlyAspGlyProAla 270
 Db 422 ATCCTGAGAAACCCCTCTGATTAATGGAATTTGGTGGCCCTACCTCTGGTGGATGGCCCTGCT 481
 QY 271 GlyAlaTyrThrArgGlnAlaGlyPheThrPalatTyrTyrGluIleCysThrPheLeuArg 290
 Db 482 GGGCCCTATACAGACAGCTGGGTTCTGGGCTACTATGAGATTTGCACCTTTCTGAGA 541
 QY 291 SerGlyAlaThrGluValTrpAspAlaSerGlnGluValProTyrAlaTyrLysAlaAsn 310
 Db 542 AGTGGAGCCACTGAGGCTGGAGTGGCTCCCAAGAGTGGCCCTATGCTTATAAGGCCAAC 601
 QY 311 GluTrpLeuGlyTyrAspAsnIleLysSerPheSerValLysAlaGlnTrpLeuLysGln 330
 Db 602 CAGTGGCTTGGCTATGACAATATCAAGAGCTTCAGTGTTAAGGCTCAGTGGCTTAAGCAG 661
 QY 331 AsnAsnPheGlyGlyAlaMetIleThrPalalaleAspLeuAspAspPhe-ThrGlySerPh 350
 Db 662 AACAAATTTGGAGGTCCTGATCTGGGCCATTTGACCTTGTATGACTTCAACTTGGCTCTTT 721
 QY 350 eCysAspGlnGly-LysPhePro-LeuThrSer-ThrLeuAsn-LysAlaLeuGlyIle 369
 Db 722 CTGTGATCAGGGAAATTTCTTCTGACTTCTAACTTTGAACAAAGCCCTGGGCAAT 781
 QY 369 eThrGluGlyCysThrAlaProAspValProSerGluProValThrThrProGlyLys 389
 Db 782 CCAGTGGAGTTGACAGT-CCTGACGTGGCTTCCGAGCCAGTGAAGTACTACTCTCTCCAGGA 840
 QY 389 er 389
 ||

```

Db      841 GT 842
RESULT 10
BQ231064
DEFINITION! BQ231064 872 bp mRNA linear EST 02-MAY-2002
IMAGE:6051490 5', mRNA sequence.
ACCESSION BQ231064
VERSION BQ231064.1 GI:20412464
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 872)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13305 row: g column: 11
High quality sequence stop: 724.
Location/Qualifiers
1. .872
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone_lib="IMAGE:6051490"
/lab_host="NCI CGAP St1"
/notes="Organ: stomach; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.77 kb. Library constructed by Life
Technologies. Note: this is a NCI CGAP Library."
BASE COUNT 218 a 222 c 205 g 225 t 2 others
ORIGIN
Alignment Scores:
Pred. No.: 4,09e-138 Length: 872
Score: 1400.00 Matches: 262
Percent Similarity: 98.87% Conservative: 0
Best Local Similarity: 98.87% Mismatches: 1
Query Match: 56.16% Indels: 2
DB: 14 Gaps: 0
US-10-004-219b-9 (1-452) x BQ231064 (1-872)
QY 1 TyrAsnLeuLeuCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlySerPhe 20
Db 68 TACAATCTGATATGCTATTTCACCAATCGGCCAGTATCGGCAGCTCTGGGAGCTTC 127
QY 21 LysProAspAspIleAsnProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet 40
Db 128 AAGCCTGATGACATTAACCCCTGCTGTACTCCTGATCTATCCCTTTGCTGGGATG 187
QY 41 GlnAsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrLysAlaPheAsn 60
Db 188 CAGAACAATGAGATCACCACCATAGATGATGATGATGATGATGATGATGATGAT 247
QY 61 AspLeuLysAsnArgAsnSerLysLeuLysThrLeuLeuAlaIleGlyLysTrpAsnPhe 80
Db 248 GACTTGAATAAAGGACAGCAACCAATGAAACCCCTCTGCAATTTGGAGCTTGAACCTTT 307
QY 81 GlyThrAlaProPheThrThrMetValSerThrSerGlnAsnArgGlnThrPheIleThr 100
Db 308 GGAACTGCTCTTCTACTACCATGGTTTCCACTTCTCAGAACCGCCAGACCTTCATTACC 367

```

```

QY 101 SerValIleLysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyr 120
Db 368 TCAGTCATCAAAATTTCTGGTCAGTATGGTTTGTGAGCTGGACCTGGAGCTGGGAATAC 427
QY 121 ProGlySerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValLysGlu 140
Db 428 CCAGGCTCACGTGGGAGCCCTCTCAGGACCAAGCATCTCTTCACTGTCTCTGGTGAAGGAA 487
QY 141 MetArgGluAlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThr 160
Db 488 ATGCGTGAAGCTTTTGGAGGAGGCTATTGAGACCAACAGGCCAGACTGATGTTACT 547
QY 161 AlaAlaValAlaGlyGlyIleSerAsnIleGlnAlaGlyTyrGluIleProGluLeuSer 180
Db 548 GTCTGTGTAGCTGGTGGGATTTCCAAACATCCAGGCTGGCTATGAGATCCCTGAACTTTCT 607
QY 181 LysTyrLeuAspPheIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr 200
Db 608 AAGTACCTGGATTTTCATCATGTCATGATGATGATGATGATGATGATGATGATGATGAT 667
QY 201 ThrGlyGluAsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeu 220
Db 668 ACTGGGAGATGATGCTCTTTTCAAAATACCTACTGAGCTGGTAGCAATGCCCTACCTC 727
QY 221 AsnValAspTyrValMetAsnTyrTrpLysAsnAsnGlyAlaProAlaGluLysLeuIle 240
Db 728 AATGTGGATTATGTCATGAACCTATTGGAAGAACAATGGAGCCCGAGCTGAGAAGCTCAT 787
QY 241 ValGlyPheProGluTyrGlyHisThr-PheIleLeuArgAsnProSerAspAsn-GlyI 260
Db 788 GTTGGATTCCCGAGAGTATGGACACACCCCTTCATCTGAGAAACCCCTCTGTAATGGGAA 847
QY 260 LeGlyAlaPro 263
Db 848 TTGGTGGCCCCC 858
RESULT 11
BQ231064
LOCUS BQ231064
DEFINITION BQ231064 872 bp mRNA linear EST 18-SEP-2001
IMAGE:6051490 5', mRNA sequence.
ACCESSION BQ231064
VERSION BQ231064.1 GI:15649582
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 923)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11899 row: n column: 08
High quality sequence stop: 795.
Location/Qualifiers
1. .923
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone_lib="IMAGE:5353927"
/sex="female, virgin"
/tissue_type="infiltrating ductal carcinoma"
/dev_stage="5 months"
FEATURES
source

```

/lab host="DH10B"

/note="organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dr.
Library constructed by Life Technologies. Investigator
providing samples: Jeffrey Green, M.D., NIH"

BASE COUNT 221 a 232 c 241 g 229 t

ORIGIN

Alignment Scores:

Pred. No.: 1,52e-137 Length: 923
Score: 1395.00 Matches: 277
Percent Similarity: 90.00% Conservativity: 2
Best Local Similarity: 89.35% Mismatches: 21
Query Match: 55.96% Indels: 11
DB: 13 Gaps: 2

US-10-004-219B-9 (1-452) x B1686954 (1-923)

QY 126 SerProGlnAspLysHisLeuPheThrValLeuValLysGluMetArgGluAlaPhe 145
DB 2 AGCCCTCTCAGCAAGCATCTCTTCACTGCTGTTGAAGAAATGCGTGAAGCTTTT 61
QY 146 GluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThrAlaAlaValalagly 165
DB 62 GAGCAGAGGCTATTGAGAGCAACAGGCCAGACTGATGTTACTGCTGCTAGCTGGT 121
QY 166 GlyIleSerAsnIleGlnAlaGlyTyrGluIleProGluLeuSerLysTyrLeuAspPhe 185
DB 122 GGGATTTCAACATCCAGCTGGCTATGAGATCCCTGAACCTTCTTAGTACCTGGATTTC 181
QY 186 IleHisValMetThrTyrAspLeuHisGlySerTyrGluGlyTyrThrGlyGluAsnSer 205
DB 182 ATCCATGTCATGACATATGACCTCCATGGCTCTCTGGAGGGCTACACTGGGAGAAATAGT 241
QY 206 ProLeuTyrGlyTyrProThrGluThrGlySerAsnAlaTyrLeuAsnValAspTyrVal 225
DB 242 CTTCTTTAAATACCTCTACTGAGACTGGTGAATGCTTACCTACCTCAATGTGATTATGTC 301
QY 226 MetAsnTyrTyrLysAsnAsnGlyAlaProAlaGluLysLeuIleValGlyPheProGlu 245
DB 302 ATGAATATTGGAAGAACATGAGCCCGAGCTGAGAGAGCTCATTTGGTGGATCCAGAG 361
QY 246 TyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIleGlyAlaProThrSer 265
DB 362 TATGGACACACCTTCTCTGAGAAACCCCTCTGATAATGGAATTTGGTCCCTACCTCT 421
QY 266 GlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTyrAlaTyrTyrGluIle 285
DB 422 GGTGATGGCCCTGCTGGGCCCTATACACAGAGGCTGGGTTCTGGGCCCTACTATGAGATT 481
QY 286 CysThrPheLeuArgSerGlyAlaThrGluValTyrAspAlaSerGlnGluValProTyr 305
DB 482 TGACCTTTCTGAGAAGTGAGCCACTGAGGTCTGGGATGCTCCCAAGAGTGGCCCTAT 541
QY 306 AlaTyrLysAlaAsnGluTyrLeuGlyTyrAspAsnIleLysSerPheSerValLysAla 325
DB 542 GCTATAGGCCCAACGAGTGGCTTGGCTATGACATATCAAGAGCTTCAGTGTAAAGCT 601
QY 326 GlnTyrLeuLysGlnAsnAsnPheGlyGlyAlaMetIleTyrAlaIleAspLeuAspAsp 345
DB 602 CAGTGGCTTAAGCAACAAATTTGAGGTGGCCATGATCTGGGCCATTGAACCTTGATGAC 661
QY 346 PheThrGlySerPheCysAspGlnGlyLys--PheProLeuThrSerThrLeuAsn--L 364
DB 662 TTCATGGCTCTTCTGTGATCCAGGAACCAATTTCTCTTGACTTCTACTTTTGAACCAA 721
QY 364 ysAlaLeuGlyIleSer-ThrGluGlyCysThrAlaProAspValProSerGluProVal 383
DB 722 AAGCCCTTGGCATATCAACACTGAAGTTGCACAGCTCCCTGACGTGCTTCCGAGCCAGTG 781
QY 384 -ThrThr-ProProGlySerGlySerGlyGlyGlySerSerGlyGlySerGlyGlys 403
DB 782 AACTACTTCTCCAGGAAGTGGAGTGGGGGCTCGAAAAGCTCCGAGGGAAGCCCTGGA 841

QY 403 erGly-----PheCysAlaAspLysAlaAspGlyLeu-TyrProValalalaspasp 419
DB 842 AGGCAGTCACTTCTCTGTCGGCCAAAGCAGATGGCTACTAGCTGTGGCAGATGAC 901
QY 420 ArgAsnAlaPheTyrGln 425
DB 902 CC-AACTGCTTTTGGCAG 918

RESULT 12

BG865815

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: L14M10820 row: h column: 05

High quality sequence stop: 781.

Location/Qualifiers

1. .830

/organism="Mus musculus"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone_lib="NCI CGAP SG2"

/lab_host="DH10B (T1 phage-resistant)"

/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site 1:

NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo

dr. Average insert size 1.3 Kb. Constructed by Life

Technologies. Note: this is a NCI CGAP Library."

BASE COUNT 199 a 198 c 223 g 210 t

ORIGIN

Alignment Scores:

Pred. No.: 1,45e-137 Length: 830

Score: 1394.50 Matches: 269

Percent Similarity: 96.76% Conservativity: 0

Best Local Similarity: 96.76% Mismatches: 5

Query Match: 55.94% Indels: 6

DB: 12 Gaps: 0

US-10-004-219B-9 (1-452) x BG865815 (1-830)

QY 149 AlaIleGluSerAsnArgProArgLeuMetValThrAlaAlaValalaglyIleSer 168

DB 3 GCTATTGAGCAACAGCCAGACTGATGTTACTGCTGTAGCTGGTGGGATTTC 62

QY 169 AsnIleGlnAlaGlyTyrGluIleProGluLeuSerLysTyrLeuAspPheIleHisVal 188

DB 63 AACATCCAGGCTGGCTATGAGATCCCTGAACCTTTCTAAGTACCTGATTCATCATGTC 122

QY 189 MetThrTyrAspLeuHisGlySerTyrGluGlyTyrThrGlyGluAsnSerProLeuTyr 208

DB 123 ATGACATATGACCTCCATGGCTCCTGGGAGGCTACACTGGGGAGAAATAGTCTCTTTAC 182

209 LysTyrProThrGluThrGlySerAsnAlaTyrLeuAsnValAspTyrValMetAsnTyr 228
 183 AAATACCTTACTGAGACTGGTACCAATGCTTACCTCAATGTGATTTATGTCATGAATAT 242
 229 TrpLysAsnGlyAlaProAlaGluLysLeuLeuValGlyPheProGluTyrGlyHis 248
 243 TGAAGAACAATGGAGCCCGAGTGAAGAGCTCATTTGTTGGATTCCAGATATGGACAC 302
 249 ThrPheLeuLeuArgAsnProSerAspAsnGlyIleGlyAlaProThrSerGlyAspGly 268
 303 ACCTTCATCTCAGAAACCCCTCTGATATGAATGGTGGCCCTACCTCTCTGTGATGGC 362
 269 ProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrpAlaTyrTyrGluLeuCysThrPhe 288
 363 CTGTCTGGCCCTATACACAGAGCGTGGTCTTGGCCCTACTATGAGATTTGACCTTT 422
 289 LeuArgSerGlyAlaThrGluValTrpAspAlaSerGlnGluValProTyrAlaTyrLys 308
 423 CTGAGAAGTGGAGCCACTGAGGTCTGGGATGCTCCCAAGAGTGGCCCTATGCTATAG 482
 309 AlaAsnGluTrpLeuGlyTyrAspAsnIleLysSerPheSerValLysAlaGlnTrpLeu 328
 483 GCAACAGTGGCTTGGCTATGACATATCAAGAGCTTCAGTGTTAAGCTCAGTGGCT 542
 329 LysGlnAsnAsnPhedGlyAlaMetIleTrpAlaIleAspLeuAspPheThrGly 348
 543 AAGCAGAACAAATTTGGAGTGCCATGATCTGGCCATTGACCTTGATGACTTCACTGGC 602
 349 SerPheCysAspGlnGly-LysPheProLeuThrSerThrLeuAsnLysAlaLeuGlyI 368
 603 TCTTTCTGTGTGATCAGGGAATAATTTCTCTGACTTCTACTTTGAACAAAGCCCTGGAT 662
 368 eSerThrGluGlyCysThrAlaProAspValProSerGluProValThrProProG 388
 663 ATCCACTGAAGTTGACAGCTCTCAGCTGCTCCGAGCAGTACTTCTCCACGG 722
 388 ySer--GlySerGlyGlySerSerGly-GlySerSerGlyGlySerGlyPheCysAl 407
 723 AAAGTGGGAGTGGGGTGGAGTCCGAGGAGAGCTCTGGAGGAGTGGATTTCTGTGC 782
 407 alaPlyAlaAspGlyLeuTyrProValAlaAspAspArgAsnAla 422
 783 CGACAAGCAGATGGCTC-TACCCTGT-GCAGATGACAGATGGCT 826
 RESULT 13
 AV072529 1007 bp mRNA linear EST 24-OCT-2001
 AV072529 Mus musculus stomach C57BL/6J adult Mus musculus cDNA
 clone 2200004B09, mRNA sequence.
 AV072529
 AV072529.2 GI:16380963
 EST.
 house mouse.
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 1007)
 Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.,
 Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J., Konno, H., Kouda,
 M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M.,
 Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki,
 D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H.,
 Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toya, T.,
 Muramatsu, M. and Hayashizaki, Y.
 RIKEN Mouse ESTs (Arakawa, T., et al. '2001)
 Unpublished (2001)
 On Jun 24, 1999 this sequence version replaced gi:5192357.
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC) Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222

Fax: 81-45-503-9216
 Email: genome-res@gsc.riken.go.jp,
 URL: http://genome.gsc.riken.go.jp/
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh
 M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
 Watahiki, M., Yoneda, Y., Iehikawa, T., Ozawa, K., Tanaka, T., Matsura
 S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and
 Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.
 10 (11), 1757-1771 (2000)
 Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara
 Y. and Hayashizaki, Y.
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanaka, I., Aizawa
 K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and
 Hayashizaki, Y.
 Computational Analysis of Full-length Mouse cDNAs Compared with
 Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
 Please visit our web site (http://genome.gsc.riken.go.jp/) for
 further details.
 cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.

FEATURES

Location/Qualifiers
 1..1007
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone_lib="2200004B09"
 /clone_lib="Mus musculus stomach C57BL/6J adult"
 /sex="male"
 /tissue_type="stomach"
 /dev_stage="adult"
 /note="Organ: mammary gland; Vector: pT73D-Pac (Pharmacia
 Ri; 1st strand cDNA was primed with a Not I - oligo(dT)
 primer [5',
 TGTACCAATCTGAAGTGGAGCGCCGCGATGTTTCTTTTCTTTTCTTTTCTTTT
 T 3']; double-stranded cDNA was ligated to Eco RI
 adaptors (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of the modified pT73 vector.
 RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
 constructed and normalized by Bento Soares and M. Fatima
 Bonaldo."

BASE COUNT 248 a 246 c 243 g 270 t

ORIGIN

Alignment Scores:

Pred. No.: 4, 77e-136 Length: 1007
 Score: 1381.50 Matches: 269
 Percent Similarity: 95.44% Conservative: 3
 Best Local Similarity: 94.39% Mismatches: 12
 Query Match: 55.42% Indels: 6
 DB: 9 Gaps: 1

US-10-004-219B-9 (1-452) x AV072529 (1-1007)

QY 168 SerAsnIleGlnAlaGlyTyrGluLeuSerLysTyrLeuAspPheIleHis 187
 Db 60 TCCAACATCCAGCTGCTATGAGAT---CCTTGACTTCTTAAGTC-CTTGATTTCATCCAT 115
 QY 188 ValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThrGlyGluAsnSerProLeu 207
 Db 116 GTCATACCATATGACTTCCATGGCTCCTGGGAGGCTACACTGGGAGAGTATGCT-CTT 174

208 TyrLysTyrProThrGluThrGlySerAsnAlaTyrLeuAsnValAspTyrValMetAsn 227
175 TATCAATACCTTACTGAGCTGTAGCAATGCTACCTCAATGTGATGATGATGATGATGAT 234
228 TyrTrpLysAsnAsnGlyAlaProAlaGluLysLeuIleValGlyPheProGluTyrGly 247
235 TATTGGAAGAACAAATGGAGCCCGAGCTGAGAGCTCAATGTTGA-TTCCAGAGATATGGA 293
248 HisThrPheIleLeuArgAsnProSerAspAnGlyIleGlyAlaProThrSerGlyAsp 267
294 CACACCTTCATCTCAGAAACCCCTCTAT-ATGGAATGT-GCCCTACCTCTGCTGAT 351
268 GlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTrpAlaTyrTyrGluIleCysThr 287
352 GGCCTCTGCTGGCCCTATATACAGAGCGTGGTTCCTGGCCCTACTATGAGATTGCGACC 411
288 PheLeuArgSerGlyAlaThrGluValTrpAspAlaSerGlnGluValProTyrAlaTyr 307
412 TTCTCTGAGAAGTGGAGCCACTGAGTCTGGGATGCTCCCAAGAAAGTGCCTATGCCTAT 471
308 LysAlaAsnGluTrpLeuGlyTyrPheAsnIleLysSerPheSerValLysAlaGlnTrp 327
472 AAGGCCAACAGTGGCTTGGCTATGACATATCAAGAGCTTCAGTGTAAAGCTCAGTGG 531
328 LeuLysGlnAsnAsnPheGlyGlyAlaMetIleTrpAlaIleAspLeuAspPheThr 347
532 CTATAGCAGACAATTTTGGAGTGGCCATGATCTGGCCATTTGACCTTGATGACTTCAT 591
348 GlySerPheCysAspGlnGlyLysPheProLeuThrSerThrLeuAsnLysAlaLeuGly 367
592 GGCCTCTTCTGTGATCAGGAAATTTCTCTGACTTCTTACTTTGAACAAAGCCCTTGGC 651
368 IleSerThrGluGlyCysThrAlaProAspValProSerGluProValThrThrProPro 387
652 ATATCACTGAAAGTGGACAGCTCTGACGTCCTTCGAGCCAGTACTCTCTCCA 711
388 GlySerGlySerGlyGlyGlySerSerGlyGlySerSerGlyGlySerGlyPheCysAla 407
712 GGAAGTGGAGTGGGGTGGAGCTCCGAGGAGACTCTGGAGGAGTGGATCTGTGTC 771
408 AspLysAlaAspGlyLeuTyrProValAlaAspAspArgAsnAlaPheTrpGlnCysIle 427
772 GACAAAGCAGATGGCTCTACCTGCTGGCAGATGACAGAAATGCTTTTGGCAGTGCATC 831
428 AsnGlyIleThrTyrGlnGlnHisCysGlnAlaGlyLeuValPheAspThrSerCysAsn 447
832 AATGGAATCACATACAGCAGCATTTGTCAGCAGGCTTGTGTTTGTATACCAAGCTGTAAT 891
448 CysCysAsnTrpPro 452
892 TGCTGCACTGGCCA 906

RG871434 857 bp mRNA linear EST 29-MAY-2001
602790636F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4921784 5',
RNA sequence.
RG871434
RG871434.1 GI:14221974
EST.
house mouse.
Mus musculus.
ORGANISM
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 857)
NIH-MGC <http://mgi.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM10840 row: h column: 09
High quality sequence stop: 782.
Location/Qualifiers
1. 857
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4921784"
/lab host="NCI CGAP SG2"
/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site 1:
Noti; Site 2: SalI; Cloned unidirectionally. Primer: Oligo
dr. Average insert size 1.3 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 222 a 203 c 214 g 218 t
ORIGIN
Alignment Scores:
Pred. No.: 3.42e-134 Length: 857
Score: 1363.00 Matches: 273
Percent Similarity: 95.50% Conservative: 3
Best Local Similarity: 94.46% Mismatches: 8
Query Match: 54.67% Indels: 8
D: 12 Gaps: 1
US-10-004-219B-9 (1-452) x BG871434 (1-857)
QY 48 IleGluTrpAsnAspValThrLeuTyrLysAlaPheAsnAspLeuLysAsnArgAsnSer 67
DB 1 ATAGAAATGAATGATGTACTCTCTATTAAGCTTCAATGACTTGAAGAAACAGGACAGC 60
QY 68 LysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPheGlyThrAlaProPheThrThr 87
DB 61 AAACCTGAAACCCCTCTCGCAATGGAGCTGGAACTTTGGAACCTGCTCTTCACTACC 120
QY 88 MetValSerThrSerGlnAsnArgGlnThrPheIleThrSerValIleLysPheLeuArg 107
DB 121 ATGGTTTCACCTCTCAGAACCGCCAGACCTTCATTACCTCAGTCATCAAAATTTCTGCT 180
QY 108 GlnTyrGlyPheAspGlyLeuAspLeuAspTrpGluTyrProGlySerArgGlySerPro 127
DB 181 CAGTATGGGTTTGTATGGAGCTGGACCTGGAGCTGGGAATACCCAGGCTCAGCTGGAGCCCT 240
QY 128 ProGlnAspLysHisLeuPheThrValLeuValLysGluMetArgGluAlaPheGluGln 147
DB 241 CCTCAGGACAGCATCTCTTCACTGCTCTGGTGAAGAAATGCGTGAAGCTTTTGAGCAG 300
QY 148 GluAlaIleGluSerAsnArgProArgLeuMetValThrAlaAlaValAlaGlyGlyIle 167
DB 301 GAGGCTATTGAGAGCAACAGAGCCAGACTGATGGTTACTGTCTGTGTAGCTGTGGGATT 360
QY 168 SerAsnIleGlnAlaGlyTyrGluIleProGluLeuSerLysTyrLeuAspPheIleHis 187
DB 361 TCCAAATCCAGGCTGGCTATGAGATCCCTGAACCTTCTTAAGTACCTGATTTTCATCAT 420
QY 188 ValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThrGlyGluAsnSerProLeu 207
DB 421 GTCATGACATATGACCTCCATGGCTCTCTGGAGGGCTACACTGGGAGAAATAGTCTCTTT 480
QY 208 TyrLysTyrProThrGluThrGlySerAsnAlaTyrLeuAsnValAspTyrValMetAsn 227
DB 481 TACAAATACCTTACTTGAGCTAGTCAATGCTACCTCAATGTGATGATGATGATGATGAT 540
QY 228 TyrTrpLysAsnAsnGlyAlaProAlaGluLysLeuIleValGlyPheProGluTyrGly 247
DB 541 TATTGGAAGAACAAATGGAGCCCGAGCTGAGAGCTCAATGTTGGAATTCACAGATATGGA 600
QY 248 HisThrPheIleLeuArgAsnProSerAspAsnGlyIleGlyAlaProThrSerGlyAsp 267


```

|||||
601 CACACCTTCATCTCGAGAAACCTCT-GATAATGGAATTGGTCCCTACCTCTCGTGAT 659
|||||
268 GlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTyrAlaTyrTyrGlu-IleCysT 287
|||||
660 GGCCCTCGCTGGGCCCTATACACAGAGGCTGGGCTCTGGGGCTACTATGAGAATTTGCA 719
|||||
287 hrPheLeuArgSer-GlyAlaThrGlu-ValTyrAspAlaSerGlnGluValProTyrAl 306
|||||
720 CCTTCTGAGAGTGGAGCCACTGAGGCTCTGGGATGA-CTCCAAGAAGTGGCCTATGC 778
|||||
306 aTyrLysAlaAsnGluTyrLeuGlyTyrAspAsnIle---LysSerPheSerValLysAl 325
|||||
779 CTATAAGGCCAAGAGTGGCTTCGATGATGACCC-ATATTCAAGAGCTCAATGTTAAGGC 837
|||||
325 aGlnTyrLeuLysGlnAsn 331
|||||
838 TCAGTGGGTTAAGCAGAAC 856

RESULT 15
LOCUS BG866853
DEFINITION 800 bp mRNA linear EST 29-MAY-2001
mRNA sequence.
ACCESSION 602786689F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4912438 5',
VERSION BG866853
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 800)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapsb@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10816 row: b column: 23
High quality sequence stop: 787.
FEATURES
    source
        1..800
            Location/Qualifiers
                organism="Mus musculus"
                strain="FVB/N"
                db_xref="taxon:10090"
                clone="IMAGE:4912438"
                /clone_lib="NCI_CGAP_SG2"
                /lab_host="DH10B (T1 phage-resistant)"
                /notes="Organ: salivary gland; Vector: pCMV-SPORT6; Site 1:
                NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo
                dt. Average insert size 1.3 kb. Constructed by Life
                Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 194 a 190 c 205 g 211 t
ORIGIN
Alignment Scores:
Pred. No.: 3 54e-133 Length: 800
Score: 1353.00 Matches: 257
Percent Similarity: 97.36% Conservatives: 1
Best Local Similarity: 96.98% Mismatches: 3
Query Match: 54.27% Indels: 4
DB: 12 Gaps: 0

US-10-004-219B-9 (1-452) x BG866853 (1-800)

QY 104 LysPheLeuArgGlnTyrGlyPheAspGlyLeuAspLeuAspTyrProGlySer 123
|||||

```

```

Db 1 AAATTTCTGCGTCAATGATGGGTTTGATGGACTGGACCTGGAGTACCGGAATACCCAGGCTCA 60
QY 124 ArgGlySerProGlnAspLysHisLeuPheThrValLeuValLysGluMetArgGlu 143
Db 61 CTTGGAGAGCCCTCTCAGAGACAGCATCTCTCACTGCTCTGGTGAAGGAATGGGTGAA 120
QY 144 AlaPheGluGlnGluAlaIleGluSerAsnArgProArgLeuMetValThrAlaAlaVal 163
Db 121 GCTTTTGAGCAGGAGGCTATTGAGAGCAACAGGCCAGCATGATGGTTACTGCTGCTGA 180
QY 164 AlaGlyGlyIleSerAsnIleGlnAlaGlyTyrGluIleProGluLeuSerLysTyrLeu 183
Db 181 GCTGTGGGATTTCCAACTCCAGGCTGGCTATGATGATCCCTGAACTTTCTAAGTACCTG 240
QY 184 AspPheIleHisValMetThrTyrAspLeuHisGlySerTyrGluGlyTyrThrGlyGlu 203
Db 241 GATTTTCATCATGTCATGACATATGACCTCCATGGCTCTGGAGGGGTACACTGGGGAG 300
QY 204 AsnSerProLeuTyrLysTyrProThrGluThrGlySerAsnAlaTyrLeuAsnValAsp 223
Db 301 AATAGTCTCTTTTCAAAATACCTACTGAGACTGGTAGCAATGCCCTACTCAATGTGAT 360
QY 224 TyrValMetAsnTyrTyrLysAsnGlnGlyAlaProAlaGluLysLeuIleValGlyPhe 243
Db 361 TATGTTCATGAACACTATTGGAAGAACAAATGAGGCCAGCTGAGAACTCATTTGTGATTC 420
QY 244 ProGluTyrGlyHisThrPheIleLeuArgAsnProSerAspAsnGlyIleGlyAla-Pr 263
Db 421 CCAGAGTATGACACACCTTCTCTCTGAGAAACCCCTCTGATATGGAATGGTGGCCCC 480
QY 263 oThrSerGlyAspGlyProAlaGlyAlaTyrThrArgGlnAlaGlyPheTyrAlaTyrTy 283
Db 481 TACCTCTGGTGATGGCCCTGCTGGGCCCTATACACAGACAGAGCTGGGTTCTGGGCCCTACTA 540
QY 283 rGluIleCysThrPheLeuArgSerGlyAlaThrGluValTyrAspAlaSerGlnGluVa 303
Db 541 TGAGATTTGCACTTTCTGAGAAGTGGAGCCCACTGAGGTCTGGGATGGCTCCCAAGAAGT 600
QY 303 lProTyrAla-TyrLysAlaAsnGluTyrLeuGlyTyrAspAsnIleLysSerPheSerV 323
Db 601 GCCCTATGCCCTATTAAGGCCAACGAGTGGCTTGGCTATGACATATCAAGAGCTTCAGTG 660
QY 323 aLysAlaGlnTyrLeuLysGlnAsnAsnPhe-GlyGlyAlaMetIleTyrAlaIleAsp 342
Db 661 TTAAGGCTCAGTGGCTTAAGCAGAACAAATTTGGAGGTCCTATGATCTGGGCCATTGAA 720
QY 343 LeuAspAspPheThrGlySerPheCysAspGlnGly-LysPheProLeuThrSerThrIle 362
Db 721 CTTGATGACTTCACTTGTGTTCTTCTGATCAGGCCAAAAATTTCTCTGACTTCTACTTT 780
QY 362 uAsnLys 364
Db 781 GAACAAG 787

```

Search completed: June 30, 2003, 01:07:36
Job time : 1505.73 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2003, 20:50:19 ; Search time 29.6228 Seconds
(without alignments)
2001.720 Million cell updates/sec

Title: US-10-004-219B-10

Perfect score: 2398

Sequence: 1 AKLVCFYTNWYQGEARF.....QQSCPTGLVFSNCKCCTWN 445

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002.*

- 1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2398	100.0	466	19	AAW40259 Human chitinase pr
2	2398	100.0	466	20	AAW42425 MO-218 clone of hu
3	2398	100.0	466	22	AAE00432 Human chitinase pr
4	2398	100.0	466	23	ABE76291 Human chitinase.
5	2394	99.8	466	18	AAW08594 Human 50 kDa chiti
6	2392	99.7	466	19	AAW40260 Human chitinase pr
7	2392	99.7	466	20	AAW42426 MO-13B clone of hu
8	2392	99.7	466	22	AAE00433 Human chitinase pr
9	2392	99.7	466	23	ABE76292 Human chitinase.
10	2391	99.7	466	18	AAW31498 Human chitotriosid

11	2372	98.9	464	18	AAW31497 Human chitotriosid
12	1982	82.7	373	19	AAW40261 Human chitinase pr
13	1982	82.7	373	20	AAW42427 Clone of the C-ter
14	1982	82.7	373	23	ABB76293 Human chitinase tr
15	1974	82.3	373	19	AAW40262 Human chitinase pr
16	1974	82.3	373	20	AAW42428 Chitinase amino ac
17	1974	82.3	373	23	ABB76294 Human chitinase an
18	1931	80.5	387	18	AAW08585 Human 39 kDa chiti
19	1333	55.2	455	22	AAW62541 Disease treatment
20	1333	55.2	476	22	AAW62543 Amino acid sequenc
21	1053.5	43.9	421	19	AAW53122 Amino acid sequenc
22	1045.5	43.6	385	19	AAW53123 Human cartilage gp
23	1045.5	43.6	385	19	AAW47033 Human cartilage gp
24	1045.5	43.6	416	19	AAW47034 Human cartilage gp
25	1045.5	43.6	421	19	AAW53121 Amino acid sequenc
26	1045.5	43.6	423	18	AAW35930 Human cartilage gl
27	1035.5	43.2	383	18	AAW26751 Bovine whey protei
28	1015.5	42.3	696	22	ABE58615 Drosophila melanog
29	1010.5	42.1	383	21	AAW03442 Gp38k protein sequ
30	981	40.9	368	23	ABB97580 Novel human protei
31	967.5	40.3	377	22	AAW62545 Disease treatment
32	967.5	40.3	398	22	AAW62544 Disease treatment
33	964.5	40.2	537	16	AAW73391 Bovine oviduct spe
34	942	39.3	718	16	AAW73392 Murine oviduct spe
35	927	38.7	668	16	AAW73393 Hamster oviduct sp
36	919.5	38.3	4498	22	ABE58595 Drosophila melanog
37	848	35.4	460	22	ABE64366 Drosophila melanog
38	824	34.4	554	18	AAW01824 Manduca sexta larv
39	824	34.4	554	21	AAW07183 Manduca sexta gut
40	797	33.2	321	9	AAW61342 Polypeptide involv
41	786.5	32.8	595	22	ABE71737 Drosophila melanog
42	762	31.8	490	21	AAW52535 D. pteronyssius 98
43	762	31.8	490	23	AAU96339 Der HMW-map polype
44	762	31.8	509	21	AAW52533 D. pteronyssius 98
45	762	31.8	509	23	AAU96337 Der HMW-map polype

ALIGNMENTS

RESULT 1	
AAW40259	
ID	AAW40259 standard; Protein; 466 AA.
XX	
AC	AAW40259;
XX	
DT	15-JUN-1998 (first entry)
XX	
DE	Human chitinase protein from clone MO-218.
XX	
KW	Chitinase; human; fungal infection; immunogen; diagnosis; treatment;
KW	Gaucher's disease; transgenic; detection; hybridisation; antifungal;
KW	rheumatoid arthritis; overexpression; extracellular matrix.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	Peptide
FT	Protein
FT	Location/Qualifiers
FT	1..21
FT	/label= signal
FT	22..466
FT	/label= chitinase
FT	/note= "from clone MO-218"
XX	
PN	WO9747752-A1.
XX	
PD	18-DEC-1997.
XX	
PF	16-JUN-1997; 97WO-US10460.
XX	
DR	14-JUN-1996; 96US-0663618.
XX	
PA	(ICOS-) ICOS CORP.
XX	

PI Gray PW;

XX WPI: 1998-052316/05.
XX N-PSDB; AAV10435.

XX Nucleic acids encoding human chitinase - useful as antifungal
XX agents, especially in combination with other antifungals

XX Claim 1; Page 41-42; 63pp; English.

XX This sequence represents a novel human chitinase isolated from clone
XX MO-218. Chitinases are useful for treating or preventing fungal infection
XX and as immunogens for generating antibodies which are used to purify,
XX detect and quantify chitinases, e.g. for diagnosis of Gaucher's disease.
XX The nucleic acid sequence of the chitinase is also useful as a probe to
XX identify and isolate genomic DNA encoding chitinases or similar proteins,
XX or cells expressing them or to generate transgenic ('knockout') rodents.
XX It can also be used in hybridisation assays and to detect genetic
XX alterations in the chitinase gene related to disease. Agents that inhibit
XX this protein may be useful in treatment of Gaucher's disease and
XX rheumatoid arthritis, where overexpression of the protein can damage
XX the extracellular matrix. Chitinase also improves the activity of other
XX antifungal agents and may allow a reduction in the dose of such agents,
XX and thus of their side effects.

XX Sequence 466 AA;

Query Match 100.0%; Score 2398; DB 19; Length 466;
Best Local Similarity 100.0%; Pred. No. 7.8e-221;
Matches 445; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKLCYFTNWAQYRQGEARFLPKDLPSLCTHLIYAFAGTNNHOLSTWENDETLQEFN 60
Db 22 AKLCYFTNWAQYRQGEARFLPKDLPSLCTHLIYAFAGTNNHOLSTWENDETLQEFN 81
Qy 61 GLKKNPKLTLAIGGNFCTQKFTDMVATANNRQTFVNSAIRFLRKYSFGDLDDWEY 120
Db 82 GLKKNPKLTLAIGGNFCTQKFTDMVATANNRQTFVNSAIRFLRKYSFGDLDDWEY 141
Qy 121 PGSQSPAVDKERTTLVQDLANAFQEAQTSKGERLLLSAAVPAGQTYDAGYEVNDKIA 180
Db 142 PGSQSPAVDKERTTLVQDLANAFQEAQTSKGERLLLSAAVPAGQTYDAGYEVNDKIA 201
Qy 181 QNLDFVNLMAVDFHGSMEKVTGHNSPLYKROESGAALSNVDAVQWQKQTPASKLI 240
Db 202 QNLDFVNLMAVDFHGSMEKVTGHNSPLYKROESGAALSNVDAVQWQKQTPASKLI 261
Qy 241 LGMTYGRSFTLASSSDTRVGAPATGSGTPGFTKEGGMLAYEYVCSWKGATKQRIQDQK 300
Db 262 LGMTYGRSFTLASSSDTRVGAPATGSGTPGFTKEGGMLAYEYVCSWKGATKQRIQDQK 321
Qy 301 VPYIFRDQWQVGFDDVESFRTKVSYLKQKGLGGAMWALDLDLDDFAGSCNQRYPILQTL 360
Db 322 VPYIFRDQWQVGFDDVESFRTKVSYLKQKGLGGAMWALDLDLDDFAGSCNQRYPILQTL 381
Qy 361 RQELSLPVLPGSTPELEVPKQPSSEPHGSPGQDTFCQKADGLVNPFRSSFYSCA 420
Db 382 RQELSLPVLPGSTPELEVPKQPSSEPHGSPGQDTFCQKADGLVNPFRSSFYSCA 441
Qy 421 AGRLFQOQSCPTGLVFSNSCKCTWN 445
Db 442 AGRLFQOQSCPTGLVFSNSCKCTWN 466

RESULT 2

AA42425
ID AA42425 standard; Protein; 466 AA.

XX AA42425;

DT 10-DEC-1999 (first entry)

XX MO-218 clone of human Chitinase, amino acid sequence.

XX KW chitin; fungal infection; immunocompromised; AIDS; chemotherapy;
XX KW organ transplant; parasite; chitin-binding; allele; vector;
XX KW truncated protein; chitin binding region.

XX OS Homo sapiens.

XX Key Location/Qualifiers
XX Peptide 1..21
XX Protein /label= Signal_peptide
XX /label= Mature_protein
XX /note= "Human Chitinase"
XX Peptide 391..445
XX /label= Chitin binding
XX /note= "Chitin binding region"

XX WO9946390-A1.

XX 16-SEP-1999.

XX 12-MAR-1999; 99WO-US05343.

XX 12-MAR-1998; 98US-0039198.

XX (ICOS-) ICOS CORP.

XX Gray PW, Tjoelker LW;

XX WPI; 1999-551417/46.

XX N-PSDB; AAZ21847.

XX Novel chitin-binding fragments of human chitinase used to treat fungal
XX infections in animals

XX Example 1; Page 58-59; 83pp; English.

XX This is the amino acid sequence of an allelic form of the human
XX chitinase enzyme, which is capable of degrading Chitin (a linear
XX homo polymer of beta-1,4-linked N-acetylglucosamine residues)
XX The sequence also includes a Chitin binding peptide region, which lacks
XX any chitinase activity, but which can be used to target anti-fungal
XX agents to chitin containing fungal cells.
XX Chitinase fragments can be used to screen for proteins or other
XX molecules that specifically bind to the chitin-binding domain of human
XX chitinase or that modulate its activity. These compounds are useful for
XX immunization, as well as for purifying chitinase. These compounds are useful for
XX detection and quantification of chitinase. Polynucleotide fragments of
XX the invention are useful as a source of probes and primers, and to
XX express the proteins recombinantly. The chitinase fragments, when
XX conjugated to antifungal compounds, are used to treat animals,
XX especially humans, infected with chitin-containing parasites such as
XX fungi. Fungal infection treated include candidiasis, aspergillosis,
XX coccidioidomycosis, blastomycosis, paracoccidioidomycosis,
XX mucormycosis, histoplasmosis, cryptococcosis, chromoblastomycosis,
XX sporotrichosis, and dermatophytosis.
XX Chitin can be degraded by the enzyme chitinase. Use of whole chitinase
XX protein for treating infections, especially fungal infections, is
XX problematic. In view of the increasing incidents of life-threatening
XX fungal infection in e.g. immunocompromised individuals, there exists a
XX need for identifying new compounds for treating fungal infection. The
XX chitin-binding fragments of the present invention provide this need.

XX Sequence 466 AA;

Query Match 100.0%; Score 2398; DB 20; Length 466;

Best Local Similarity 100.0%; Pred. No. 7.8e-221;

Matches 445; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKLCYFTNWAQYRQGEARFLPKDLPSLCTHLIYAFAGTNNHOLSTWENDETLQEFN 60

Db 22 AKLCYFTNWAQYRQGEARFLPKDLPSLCTHLIYAFAGTNNHOLSTWENDETLQEFN 81

QY 61 GLKONPKLKTLLAIGGNWFGTKFTDMVATANNRQTFVNSAIRFLRYSPDGLDLDWEY 120
 DB 82 GLKONPKLKTLLAIGGNWFGTKFTDMVATANNRQTFVNSAIRFLRYSPDGLDLDWEY 141
 QY 121 PGSQSPAVDKERFTTLVQDLANAFQOEATQSGKERLLLSAAVPAGQTVVDAGYVDKIA 180
 DB 142 PGSQSPAVDKERFTTLVQDLANAFQOEATQSGKERLLLSAAVPAGQTVVDAGYVDKIA 201
 QY 181 QNLDFVNLMAVDYFHGSWEKVTGHNPSPLYKROESGAAASLNVDAAVQWLQKGTTPASKLI 240
 DB 202 QNLDFVNLMAVDYFHGSWEKVTGHNPSPLYKROESGAAASLNVDAAVQWLQKGTTPASKLI 261
 QY 241 LGMPTYGRSFTLLASSDTRVGAPATGSGTPGFTKEGGMLAYYEVCSWKATQRIQDOK 300
 DB 262 LGMPTYGRSFTLLASSDTRVGAPATGSGTPGFTKEGGMLAYYEVCSWKATQRIQDOK 321
 QY 301 VPYIFRDNQWVGFDVSEFKTKVSYLKQKGLGGAMVWALDLDLDFAGFSCNQGRYPLIOTL 360
 DB 322 VPYIFRDNQWVGFDVSEFKTKVSYLKQKGLGGAMVWALDLDLDFAGFSCNQGRYPLIOTL 381
 QY 361 ROELSLPYLPSTPELEVPKQPSPEHGPSPGQDTFCQKADGLYPNPRRSSFYSKA 420
 DB 382 ROELSLPYLPSTPELEVPKQPSPEHGPSPGQDTFCQKADGLYPNPRRSSFYSKA 441

RESULT 3

AAE00432
 ID AAE00432 standard; Protein; 466 AA.
 AC AAE00432;
 XX
 DT 19-JUN-2001 (first entry)
 XX
 DE Human chitinase protein from clone pMO-218.
 XX
 KW Human; antifungal; chitinase; immunoglobulin; Ig; therapy;
 KW fungal infection; candidiasis; aspergilliosis; coccidioidomycosis;
 KW blastomycosis; paracoccidioidomycosis; histoplasmosis; cryptococcosis;
 KW chromoblastomycosis; sporotrichosis; mucormycosis; dermatophytoses;
 KW clone pMO-218.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..21
 FT /label= Signal_peptide
 FT Protein 22..466
 FT /label= Human_mature_chitinase_protein
 FT Domain 418..466
 FT /label= Chitin_binding_domain
 FT Region 1..373
 FT /notes= "Region with triacetylchitotriose hydrolysing activity"
 FT
 FT
 XX WO200123430-A2.
 XX
 XX 05-APR-2001.
 XX
 PF 28-SEP-2000; 2000WO-US26960.
 XX
 PR 30-SEP-1999; 99US-0409918.
 XX
 XX (ICOS-) ICOS CORP.
 PA
 XX Allison DS, Dietsch GN, Gray PW, Shaw KD, Steiner BH;
 XX
 DR WPI; 2001-266141/27.
 DR N-PSDB; AAD03759.
 XX

PT Novel chitinase immunoglobulin fusion product, useful for treating
 PT fungal infections and reducing the amount of a non-chitinase antifungal
 PT agent needed for the treatment -
 XX
 PS Claim 1; Page 32-33; 39pp; English.
 XX
 CC The present invention relates to a chitinase immunoglobulin (Ig) fusion
 CC product, comprising a human chitinase fused to at least a portion of an
 CC immunoglobulin chain. The fusion product is useful for treating fungal
 CC infections (mycoses) such as candidiasis, aspergilliosis, blastomycosis,
 CC coccidioidomycosis, paracoccidioidomycosis, histoplasmosis, mucormycosis,
 CC cryptococcosis, chromoblastomycosis, sporotrichosis and dermatomycoses.
 CC The fusion protein is useful for reducing the amount of non-chitinase
 CC antifungal agent needed to exert an antifungal activity. The fusion
 CC protein is also useful for preparing
 CC a medicament for the prophylactic or therapeutic treatment of fungal
 CC infections. Chitinase immunoglobulin fusion product has unexpectedly
 CC improved serum half-life and formulation properties.
 CC The present sequence is human chitinase protein from clone pMO-218.
 CC Chitinase enzyme degrades chitin which is a homopolymer of
 CC beta-(1,4)-linked N-acetylglucosamine residues.
 XX
 SQ Sequence 466 AA;
 Query Match 100.0%; Score 2398; DB 22; Length 466;
 Best Local Similarity 100.0%; Pred. No. 7.8e-221;
 Matches 445; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AKLVCFYTNWQYRQGEARFLPKDLDPISLCTHLIYAFAGMTHQLSTTEWNETLYQEFN 60
 DB 22 AKLVCFYTNWQYRQGEARFLPKDLDPISLCTHLIYAFAGMTHQLSTTEWNETLYQEFN 81
 QY 61 GLKONPKLKTLLAIGGNWFGTKFTDMVATANNRQTFVNSAIRFLRYSPDGLDLDWEY 120
 DB 82 GLKONPKLKTLLAIGGNWFGTKFTDMVATANNRQTFVNSAIRFLRYSPDGLDLDWEY 141
 QY 121 PGSQSPAVDKERFTTLVQDLANAFQOEATQSGKERLLLSAAVPAGQTVVDAGYVDKIA 180
 DB 142 PGSQSPAVDKERFTTLVQDLANAFQOEATQSGKERLLLSAAVPAGQTVVDAGYVDKIA 201
 QY 181 QNLDFVNLMAVDYFHGSWEKVTGHNPSPLYKROESGAAASLNVDAAVQWLQKGTTPASKLI 240
 DB 202 QNLDFVNLMAVDYFHGSWEKVTGHNPSPLYKROESGAAASLNVDAAVQWLQKGTTPASKLI 261
 QY 241 LGMPTYGRSFTLLASSDTRVGAPATGSGTPGFTKEGGMLAYYEVCSWKATQRIQDOK 300
 DB 262 LGMPTYGRSFTLLASSDTRVGAPATGSGTPGFTKEGGMLAYYEVCSWKATQRIQDOK 321
 QY 301 VPYIFRDNQWVGFDVSEFKTKVSYLKQKGLGGAMVWALDLDLDFAGFSCNQGRYPLIOTL 360
 DB 322 VPYIFRDNQWVGFDVSEFKTKVSYLKQKGLGGAMVWALDLDLDFAGFSCNQGRYPLIOTL 381
 QY 361 ROELSLPYLPSTPELEVPKQPSPEHGPSPGQDTFCQKADGLYPNPRRSSFYSKA 420
 DB 382 ROELSLPYLPSTPELEVPKQPSPEHGPSPGQDTFCQKADGLYPNPRRSSFYSKA 441
 QY 421 AGRLFQSQSCPTGLVFSNSCKCCTWN 445
 DB 442 AGRLFQSQSCPTGLVFSNSCKCCTWN 466
 RESULT 4
 ABB76291
 ID ABB76291 standard; Protein; 466 AA.
 XX
 AC ABB76291;
 XX
 DT 12-AUG-2002 (first entry)
 XX
 DE Human chitinase.
 XX
 KW Chitinase; enzyme; human; fungicide; antifungal; infection;
 KW candidiasis; aspergilliosis; coccidioidomycosis; blastomycosis;

KW paracoccidioidomycosis; histoplasmosis; cryptococcosis;
 KW chromoblastomycosis; sporotrichosis; mucormycosis; dermatophytosis;
 KW Pneumocystis.
 XX Homo sapiens.

XX Key Location/Qualifiers
 FT Peptide 1..21
 FT /label= Signal_peptide
 FT Protein 22..466
 FT /label= Mature_protein

XX US6372212-B1.

XX 16-APR-2002.

XX 16-JUN-1997; 97US-0877599.

XX 14-JUN-1996; 96US-0663618.

XX (ICOS-) ICOS CORP.

XX Gray PW;

XX WPI; 2002-442449/47.

XX N-PSDB; ABL57380.

XX Co-administering chitinase to improve the effectiveness of fungicidal
 PT drugs e.g. amphotericin B or itraconazole, useful for treating fungal
 PT infections e.g. candidiasis, coccidioidomycosis and blastomycosis -

XX Claim 1; Column 27-30; 26pp; English.

XX The present sequence is the protein sequence of human chitinase
 CC as predicted from isolated cDNA clone MO-218 (see ABL57380). It
 CC differs in only 1 amino acid residue from the chitinase sequence
 CC (see ABB76292) deduced from a second cDNA clone, having glycine
 CC at position 81 of the mature protein. Northern blots showed
 CC highest chitinase gene expression in lung and ovary tissues.
 CC Expression in lung is consistent with a protective role against
 CC pathogenic organisms that contain chitin. The invention provides
 CC human chitinase polynucleotides and polypeptides, and materials and
 CC methods for the recombinant production of human chitinase products,
 CC which are expected to be useful as products for treating fungal
 CC infections or for the development of such products. Human
 CC chitinase has a synergistic effect on the actions of other
 CC fungicides. It can be administered to improve the antifungal
 CC activity of a non-chitinase antifungal agent, especially
 CC amphotericin B or itraconazole, in the treatment of a fungal
 CC infection such as candidiasis, aspergillosis, coccidioidomycosis,
 CC blastomycosis, paracoccidioidomycosis, histoplasmosis,
 CC cryptococcosis, chromoblastomycosis, sporotrichosis, mucormycosis,
 CC dermatophytoses and Pneumocystis infections (all claimed). In
 CC particular, the fungal infection involves *Candida*, *Aspergillus*
 CC and/or *Cryptococcus* spp., whose growth is not effectively
 CC inhibited by contact with human chitinase alone.

XX Sequence 466 AA;

Query Match 100.0%; Score 2398; DB 23; Length 466;
 Best Local Similarity 100.0%; Pred. No. 7,8e-221;
 Matches 445; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKLVCFYTNWAOYRQGEARLPKDLPSLCTHLYIYAFAGMTNHLSTWENDETLQEFN 60
 DB 22 AKLVCFYTNWAOYRQGEARLPKDLPSLCTHLYIYAFAGMTNHLSTWENDETLQEFN 81
 QY 61 GLKKMNPKLTLAIGWNEFGTKFTDMVATANNROTFFNSAIRFLRKYSFDGLDLWEY 120
 DB 82 GLKKMNPKLTLAIGWNEFGTKFTDMVATANNROTFFNSAIRFLRKYSFDGLDLWEY 141
 QY 121 PGSGQSPAVDKERFTTLVQDLANAFQEAQTSKERLLLSAAVPAQGTYYVDAGYEVDKIA 180

Db 142 PGSGQSPAVDKERFTTLVQDLANAFQEAQTSKERLLLSAAVPAQGTYYVDAGYEVDKIA 201
 QY 181 QNLDFVNLWAYDFHGSWEKVTGHNSPLYKROESGAAASLNVDAAVQOVLQKGTTPASKLI 240
 Db 202 QNLDFVNLWAYDFHGSWEKVTGHNSPLYKROESGAAASLNVDAAVQOVLQKGTTPASKLI 261
 QY 241 LGMPTYGRSFTLLASSSDTRVGAPATGSGTGPFTTKEGMLAYEYVCSWKGATKQRIODQK 300
 Db 262 LGMPTYGRSFTLLASSSDTRVGAPATGSGTGPFTTKEGMLAYEYVCSWKGATKQRIODQK 321
 QY 301 VPIYIFRDQWVGFDVSEPKTKVSYLKQKLGAMVWALDLDLDDFAGFSCNQGRIPIQTL 360
 Db 322 VPIYIFRDQWVGFDVSEPKTKVSYLKQKLGAMVWALDLDLDDFAGFSCNQGRIPIQTL 381
 QY 361 ROELSILPLPSGTPELEVPKGPSPPEHGPSPGQDTFCQKADGLYPNPRRSSFYSICA 420
 Db 382 ROELSILPLPSGTPELEVPKGPSPPEHGPSPGQDTFCQKADGLYPNPRRSSFYSICA 441
 QY 421 AGRLFQOQSCPTGLVFSNSCKCCTWN 445
 Db 442 AGRLFQOQSCPTGLVFSNSCKCCTWN 466

RESULT 5

AAW08584

ID AAW08584 standard; Protein; 466 AA.

XX AAW08584;

AC AAW08584;

XX 24-MAR-1997 (first entry)

XX Human 50 kDa chitinase.

XX Chitinase; chitotriosidase; chitin; infectious disease;

XX gene therapy; vaccine; diagnosis; lysosomal lipidosis;

XX Gaucher disease; leishmaniasis; sarcoidosis; multiple sclerosis;

XX X-linked adrenoleukodystrophy; drug delivery; cosmetics; food.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..21

FT /label= Sig_peptide

FT Domain 178..198

FT /label= Catalytic_domain

XX WO9640940-A2.

XX 19-DEC-1996.

XX 06-JUN-1996; 96WO-NL00225.

XX 07-JUN-1995; 95US-0486839.

XX (UNAM) UNIV AMSTERDAM.

XX Aerts JMEG;

XX WPI; 1997-118698/11.

XX N-PSDB; AAT50833.

XX New human chitinase - used to treat or prevent infection by

XX chitin-contg. pathogens, in diagnosis and as additives to cosmetics,

XX foods, implants etc.

XX Claim 1; Page 40-42; 58pp; English.

XX Human chitinases of 50 kDa (AAW08584) and 39 kDa (AAW08585) are stable

XX to many proteases, active at pH 3-8 and up to 50 deg, and stable in

XX the circulation. They are the result of alternative splicing of

XX RNA, and can be produced on a large-scale in transformed host cells

XX using cDNA clones (see also AAT50833-34) obtd. from a human macrophage

XX library. The chitinases, optionally expressed from a gene, therapy

CC vector, are used to treat or prevent infection by chitin-contg.
CC pathogens (e.g. fungi, protozoa, helminths) and more generally to
CC degrade chitin. They are also useful in controlled-release drug
CC delivery, in cosmetics, foods and dental products, for antibody
CC prodn. and for diagnosis of diseases associated with elevated
CC chitinase levels.

XX SQ Sequence 466 AA;

Query Match 99.8%; Score 2394; DB 18; Length 466;
Best Local Similarity 100.0%; Pred. No. 1.9e-220;
Matches 444; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KLVCFYTNWAQYRQGEARFLPKDLPDLCTHLYIYAFAGMTNHQSLSTTEWDETLYQEFNG 61
DB 23 KLVCFYTNWAQYRQGEARFLPKDLPDLCTHLYIYAFAGMTNHQSLSTTEWDETLYQEFNG 82
QY 62 LKQKNPKLTLAIGGWNFGTKFTDMVATANNRQTFVNSAIRFLKYSFDGLDLDWEYP 121
DB 83 LKQKNPKLTLAIGGWNFGTKFTDMVATANNRQTFVNSAIRFLKYSFDGLDLDWEYP 142
QY 122 GSQGSFPAVDKERFTTLVQDLANAFQEAQTSKGERLLLSAAVPAGQTYVDAGYVDKIAQ 181
DB 143 GSQGSFPAVDKERFTTLVQDLANAFQEAQTSKGERLLLSAAVPAGQTYVDAGYVDKIAQ 202
QY 182 NLDFVNLMAFYDHFSGWEKVTGHNSPLYKROESGAAASLNVDAAVQQWLQKGTTPASKLIL 241
DB 203 NLDFVNLMAFYDHFSGWEKVTGHNSPLYKROESGAAASLNVDAAVQQWLQKGTTPASKLIL 262
QY 242 GMPTYGRSFTLASSSDTRVGAATSGTGPFTKEGMLAYEVCWKGATKQRTQDKV 301
DB 263 GMPTYGRSFTLASSSDTRVGAATSGTGPFTKEGMLAYEVCWKGATKQRTQDKV 322
QY 302 PYIFRDQNWGFDDVESFKTKVSYLKQGLGAMVWALDLDFFAGSCNQGRIPLIQTLR 361
DB 323 PYIFRDQNWGFDDVESFKTKVSYLKQGLGAMVWALDLDFFAGSCNQGRIPLIQTLR 382
QY 362 QELSPLYPSPGTPLEVPKPGPSEPEHGFSPGQDTFCQKADGLYPNPRRSPFSYSCAA 421
DB 383 QELSPLYPSPGTPLEVPKPGPSEPEHGFSPGQDTFCQKADGLYPNPRRSPFSYSCAA 442
QY 422 GRLFOQSCPTGLVFSNSCKCCTWN 445
DB 443 GRLFOQSCPTGLVFSNSCKCCTWN 466

RESULT 6
AAW40260
ID AAW40260 standard; Protein; 466 AA.

XX AC AAW40260;
XX DT 15-JUN-1998 (first entry)

XX DE Human chitinase protein from clone MO-13B.
XX KW Chitinase; human; fungal infection; immunogen; diagnosis; treatment;
XX KW Gaucher's disease; transgenic; detection; hybridisation; antifungal;
XX KW rheumatoid arthritis; overexpression; extracellular matrix.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers
XX FT Peptide 1..21
XX FT Protein /label= signal
XX FT /label= chitinase
XX FT /note= "from clone MO-13B"

XX PN W09747752-A1.

XX PD 18-DEC-1997.

XX

PF 16-JUN-1997; 97WO-US10460.
XX 14-JUN-1996; 96US-0663618.
XX (ICOS-) ICOS CORP.
XX PI Gray PW;
XX WPI; 1998-052316/05.
DR N-PSDB; AAV10436.
XX Nucleic acids encoding human chitinase - useful as antifungal
PT agents, especially in combination with other antifungals
PS Claim 7; Page 44-45; 63pp; English.
XX This sequence represents a novel human chitinase isolated from clone
CC MO-13B. Chitinases are useful for treating or preventing fungal infection
CC and as immunogens for generating antibodies which are used to purify,
CC detect and quantify chitinases, e.g. for diagnosis of Gaucher's disease.
CC The nucleic acid sequence of the chitinase is also useful as a probe to
CC identify and isolate genomic DNA encoding chitinases or similar proteins,
CC or cells expressing them or to generate transgenic ('knockout') rodents.
CC It can also be used in hybridisation assays and to detect genetic
CC alterations in the chitinase gene related to disease. Agents that inhibit
CC this protein may be useful in treatment of Gaucher's disease and
CC rheumatoid arthritis, where overexpression of the protein can damage
CC the extracellular matrix. Chitinase also improves the activity of other
CC antifungal agents and may allow a reduction in the dose of such agents,
CC and thus of their side effects.

XX SQ Sequence 466 AA;

Query Match 99.7%; Score 2392; DB 19; Length 466;
Best Local Similarity 99.8%; Pred. No. 2.9e-220;
Matches 444; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AKLVCFYTNWAQYRQGEARFLPKDLPDLCTHLYIYAFAGMTNHQSLSTTEWDETLYQEFN 60
DB 22 AKLVCFYTNWAQYRQGEARFLPKDLPDLCTHLYIYAFAGMTNHQSLSTTEWDETLYQEFN 81
QY 61 GLKQKNPKLTLAIGGWNFGTKFTDMVATANNRQTFVNSAIRFLKYSFDGLDLDWEY 120
DB 82 GLKQKNPKLTLAIGGWNFGTKFTDMVATANNRQTFVNSAIRFLKYSFDGLDLDWEY 141
QY 121 PGSQGSFPAVDKERFTTLVQDLANAFQEAQTSKGERLLLSAAVPAGQTYVDAGYVDKIA 180
DB 142 PGSQGSFPAVDKERFTTLVQDLANAFQEAQTSKGERLLLSAAVPAGQTYVDAGYVDKIA 201
QY 181 QNLDFVNLMAFYDHFSGWEKVTGHNSPLYKROESGAAASLNVDAAVQQWLQKGTTPASKLI 240
DB 202 QNLDFVNLMAFYDHFSGWEKVTGHNSPLYKROESGAAASLNVDAAVQQWLQKGTTPASKLI 261
QY 241 LGMPTYGRSFTLASSSDTRVGAATSGTGPFTKEGMLAYEVCWKGATKQRTQDKQ 300
DB 262 LGMPTYGRSFTLASSSDTRVGAATSGTGPFTKEGMLAYEVCWKGATKQRTQDKQ 321
QY 301 VPIYIFRDQNWGFDDVESFKTKVSYLKQGLGAMVWALDLDFFAGSCNQGRIPLIQT 360
DB 322 VPIYIFRDQNWGFDDVESFKTKVSYLKQGLGAMVWALDLDFFAGSCNQGRIPLIQT 381
QY 361 ROELSPLYPSPGTPLEVPKPGPSEPEHGFSPGQDTFCQKADGLYPNPRRSPFSYSCA 420
DB 382 ROELSPLYPSPGTPLEVPKPGPSEPEHGFSPGQDTFCQKADGLYPNPRRSPFSYSCA 441
QY 421 AGRLFQOQSCPTGLVFSNSCKCCTWN 445
DB 442 AGRLFQOQSCPTGLVFSNSCKCCTWN 466

RESULT 7
AAAY42426
ID AAY42426 standard; Protein; 466 AA.

XX AC AAY42426;
 XX DT 10-DEC-1999 (first entry)
 XX DE MO-13B clone of human Chitinase, amino acid sequence.
 XX KW chitin; fungal infection; immunocompromised; AIDS; chemotherapy;
 KW organ transplant; parasite; chitin-binding; allele; vector;
 KW truncated protein; chitin binding region.
 XX OS Homo sapiens.
 XX FH Key
 FT Peptide 1..21
 FT Protein /label= Signal_peptide
 FT /label= Mature_protein
 XX PN WO9946390-A1.
 XX PD 16-SEP-1999.
 XX PF 12-MAR-1999; 99WO-US05343.
 XX PR 12-MAR-1998; 98US-0039198.
 XX PA (ICOS-) ICOS CORP.
 XX PI Gray PW, Tjoelker LW;
 XX WPI; 1999-551417/46.
 XX N-PSDB; AAZ21848.
 XX Novel chitin-binding fragments of human chitinase used to treat fungal
 XX infections in animals -
 XX Example 1; Page 62-64; 83pp; English.
 XX This is the amino acid sequence of an allelic form of the human
 CC chitinase enzyme, which is capable of degrading Chitin (a linear
 CC homo polymer of beta-1,4-linked N-acetylglucosamine residues).
 CC The sequence also includes a Chitin binding peptide region, which lacks
 CC any chitinase activity, but which can be used to target anti-fungal
 CC agents to chitin containing fungal cells.
 CC Chitinase fragments can be used to screen for proteins or other
 CC molecules that specifically bind to the chitin-binding domain of human
 CC chitinase or that modulate its activity. These compounds are useful for
 CC immunization, as well as for purifying chitinase, as well as for
 CC detection and quantification of chitinase. Polynucleotide fragments of
 CC the invention are useful as a source of probes and primers, and to
 CC express the proteins recombinantly. The chitinase fragments, when
 CC conjugated to antifungal compounds, are used to treat animals,
 CC especially humans, infected with chitin-containing parasites such as
 CC fungi. Fungal infection treated include candidiasis, aspergillosis,
 CC coccidioidomycosis, blastomycosis, paracoccidioidomycosis,
 CC mucormycosis, histoplasmosis, cryptococcosis, chromoblastomycosis,
 CC sporotrichosis, and dermatophytoses.
 CC Chitin can be degraded by the enzyme chitinase. Use of whole chitinase
 CC protein for treating infections, especially fungal infections, is
 CC problematic. In view of the increasing incidents of life-threatening
 CC fungal infection in e.g. immunocompromised individuals, there exists a
 CC need for identifying new compounds for treating fungal infection. The
 CC chitin-binding fragments of the present invention provide this need.
 XX SQ Sequence 466 AA;
 XX Query Match 99.7%; Score 2392; DB 20; Length 466;
 XX Best Local Similarity 99.8%; Pred. No. 2.9e-220;
 XX Matches 444; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 XX 1 AKLVCFYTNNAQYRQGEARFLPKDLPDLCTHLIIYAFAGMTNHQSLSTTEWDETLYQEFN 60
 |||

DB 22 AKLVCFYTNNAQYRQGEARFLPKDLPDLCTHLIIYAFAGMTNHQSLSTTEWDETLYQEFN 81
 QY 61 GLKKNPKLKTLLAIGGWNFGTKFTDMVATANNRQTTFVNSAIRFLRKYSPDGLDLDWEY 120
 DB 82 GLKKNPKLKTLLAIGGWNFGTKFTDMVATANNRQTTFVNSAIRFLRKYSPDGLDLDWEY 141
 QY 121 PGSQSPAVDKERFTTLVQDLANAFQQAQTSKGKRLLLSAAVPAGQTVVDAGYEVDKIA 180
 DB 142 PGSQSPAVDKERFTTLVQDLANAFQQAQTSKGKRLLLSAAVPAGQTVVDAGYEVDKIA 201
 QY 181 QNLDFNLMAVDYFHGSWEKVTGHSPLYKROESGAASLNVDAAVQOVLQKGTASKLI 240
 DB 202 QNLDFNLMAVDYFHGSWEKVTGHSPLYKROESGAASLNVDAAVQOVLQKGTASKLI 261
 QY 241 LGMPTVGRSFTLLASSSDTRVGAPATGSGTPGFTKEGGMLAYEYVCSWKATQRIQDOK 300
 DB 262 LGMPTVGRSFTLLASSSDTRVGAPATGSGTPGFTKEGGMLAYEYVCSWKATQRIQDOK 321
 QY 301 VPIYFRDNQWGFDDVESFKTKVSYLKQKLGAMVWALDLDLDDFAGFSCNQGRIPLIOTL 360
 DB 322 VPIYFRDNQWGFDDVESFKTKVSYLKQKLGAMVWALDLDLDDFAGFSCNQGRIPLIOTL 381
 QY 361 ROELSLPYLPSGTPELEVPKQGPSEPHGSPGODTFCQKADGLYNPNRERSFYSCA 420
 DB 382 ROELSLPYLPSGTPELEVPKQGPSEPHGSPGODTFCQKADGLYNPNRERSFYSCA 441
 QY 421 AGRLFQOQSCPTGLVFSNCKCCTWN 445
 DB 442 AGRLFQOQSCPTGLVFSNCKCCTWN 466
 RESULT 8
 ID AAE00433
 XX AAE00433 standard; Protein; 466 AA.
 AC AAE00433;
 XX
 DT 19-JUN-2001 (first entry)
 XX Human chitinase protein from clone pMO-13B.
 DE Human chitinase protein from clone pMO-13B.
 XX Human; antifungal; chitinase; immunoglobulin; Ig; therapy;
 KW fungal infection; candidiasis; aspergillosis; coccidioidomycosis;
 KW blastomycosis; paracoccidioidomycosis; histoplasmosis; cryptococcosis;
 KW chromoblastomycosis; sporotrichosis; mucormycosis; dermatophytoses;
 KW clone pMO-13B.
 XX Homo sapiens.
 OS
 FH Key
 FT Peptide 1..21
 FT Protein /label= Signal_peptide
 FT /label= Human_mature_chitinase_protein
 FT Domain 418..466
 FT /label= Chitin_binding_domain
 FT Region 1..373
 FT /note= "Region with triacetylechitotriose
 FT hydrolysing activity"
 XX WO2001213430-A2.
 PN
 XX 05-APR-2001.
 PD
 XX 28-SEP-2000; 2000WO-US26960.
 PF
 XX 30-SEP-1999; 99US-0409918.
 PR
 XX (ICOS-) ICOS CORP.
 PA
 XX Allison DS, Dietech GN, Gray PW, Shaw KD, Steiner BH;
 PI WPI; 2001-266141/27.
 XX

DR N-PSDB; AAD03760.
XX Novel chitinase immunoglobulin fusion product, useful for treating
PT fungal infections and reducing the amount of a non-chitinase antifungal
PT agent needed for the treatment -
XX
XX Claim 1; Page 36-38; 39pp; English.
XX
XX The present invention relates to a chitinase immunoglobulin (Ig) fusion
CC product, comprising a human chitinase fused to at least a portion of an
CC immunoglobulin chain. The fusion product is useful for treating fungal
CC infections (mycoses) such as candidiasis, aspergillosis, blastomycosis,
CC coccidioidomycosis, paracoccidioidomycosis, histoplasmosis, mucormycosis,
CC cryptococcosis, chromoblastomycosis, sporotrichosis and dermatophytoses.
CC The fusion protein is useful for reducing the amount of non-chitinase
CC antifungal agent needed to exert an antifungal activity. The fusion
CC protein is also useful for preparing
CC a medicament for the prophylactic or therapeutic treatment of fungal
CC infections. Chitinase immunoglobulin fusion product has unexpectedly
CC improved serum half-life and formulation properties.
CC The present sequence is human chitinase protein from clone pMO-13B.
CC Chitinase enzyme degrades chitin which is a homopolymer of
CC beta-(1,4)-linked N-acetylglucosamine residues.
XX
SQ Sequence 466 AA;

Query Match 99.7%; Score 2392; DB 22; Length 466;
Best Local Similarity 99.8%; Pred. No. 2.9e-220;
Matches 444; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 AKLVCFYFTNWAQYRQGEARFLPKDLPSCSLTHLIYAFAGMTNHQLSTTEWDETLYQSFN 60
DB 22 AKLVCFYFTNWAQYRQGEARFLPKDLPSCSLTHLIYAFAGMTNHQLSTTEWDETLYQSFN 81
QY 61 GLKXNPKLKTLLAIGGWNFGTQKFTDMVATANNRQTFVNSAIRFLKYSFDGLDWEY 120
DB 82 GLKXNPKLKTLLAIGGWNFGTQKFTDMVATANNRQTFVNSAIRFLKYSFDGLDWEY 141
QY 121 PGSQSPAVDKERFTLLVDLANAFQQAQTSKGERLLLSAAVPAQTVVDAGYEVDKIA 180
DB 142 PGSQSPAVDKERFTLLVDLANAFQQAQTSKGERLLLSAAVPAQTVVDAGYEVDKIA 201
QY 181 QNLDFVNLMAVDFHGSWEKVTGHSNPLYKROESGAASLNVDAAVQWLQKGTTPASKLI 240
DB 202 QNLDFVNLMAVDFHGSWEKVTGHSNPLYKROESGAASLNVDAAVQWLQKGTTPASKLI 261
QY 241 LGMPYGRSFTLASSDTRVGAPATGSGTPGFTKEGGLAYEYVCSWKATKQRIQDK 300
DB 262 LGMPYGRSFTLASSDTRVGAPATGSGTPGFTKEGGLAYEYVCSWKATKQRIQDK 321
QY 301 VPIYFRDQWVGFDVGFVKTVSYLKQGLGAMVWALDLDFFAGFCNQGRIYPLIOTL 360
DB 322 VPIYFRDQWVGFDVGFVKTVSYLKQGLGAMVWALDLDFFAGFCNQGRIYPLIOTL 381
QY 361 ROELSPLPSTGTPLEVPKQPSPEHGPSPGDDTFCQKADGLYNPNRPSRFSYSCA 420
DB 382 ROELSPLPSTGTPLEVPKQPSPEHGPSPGDDTFCQKADGLYNPNRPSRFSYSCA 441
QY 421 AGRLLFQSCPTGLVFSNCKCCTWN 445
DB 442 AGRLLFQSCPTGLVFSNCKCCTWN 466

RESULT 9
AB76292
ID AB76292 standard; Protein; 466 AA.
XX
AC AB76292;
XX
DT 12-AUG-2002 (first entry)
XX
DE Human chitinase.
XX

KW Chitinase; enzyme; human; fungicide; antifungal; infection;
KW paracoccidioidomycosis; coccidioidomycosis; blastomycosis;
KW chromoblastomycosis; sporotrichosis; mucormycosis; dermatophytosis;
KW Pneumocystis.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FT Peptide 1-21
FT Protein /label= Signal_peptide
FT /label= Mature_protein
XX US6372212-B1.
XX 16-APR-2002.
XX 16-JUN-1997; 97US-0877599.
XX 14-JUN-1996; 96US-0663618.
XX (ICOS-) ICOS CORP.
XX Gray PW;
XX WPI; 2002-442449/47.
XX N-PSDB; ABL57381.

Co-administering chitinase to improve the effectiveness of fungicidal
drugs e.g. amphotericin B or itraconazole, useful for treating fungal
infections e.g. candidiasis, coccidioidomycosis and blastomycosis -
Example 1; Column 33-36; 26pp; English.

The present sequence is the protein sequence of human chitinase
as predicted from isolated cDNA clone MO-13B (see ABL57378). It
differs in only 1 amino acid residue from the chitinase sequence
(see ABB76291) deduced from a second cDNA clone, having serine
at position 81 of the mature protein. Northern blots showed
highest chitinase gene expression in lung and ovary tissues.
Expression in lung is consistent with a protective role against
pathogenic organisms that contain chitin. The invention provides
human chitinase polynucleotides and polypeptides, and materials and
methods for the recombinant production of human chitinase products,
which are expected to be useful as products for treating fungal
infections or for the development of such products. Human
chitinase has a synergistic effect on the actions of other
fungicides. It can be administered to improve the antifungal
activity of a non-chitinase antifungal agent, especially
amphotericin B or itraconazole, in the treatment of a fungal
infection such as candidiasis, aspergillosis, coccidioidomycosis,
blastomycosis, paracoccidioidomycosis, histoplasmosis,
cryptococcosis, chromoblastomycosis, sporotrichosis, mucormycosis,
dermatophytoses and Pneumocystis infections (all claimed). In
particular, the fungal infection involves Candida, Aspergillus
and/or Cryptococcus spp., whose growth is not effectively
inhibited by contact with human chitinase alone.

Sequence 466 AA;

Query Match 99.7%; Score 2392; DB 23; Length 466;
Best Local Similarity 99.8%; Pred. No. 2.9e-220;
Matches 444; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AKLVCFYFTNWAQYRQGEARFLPKDLPSCSLTHLIYAFAGMTNHQLSTTEWDETLYQSFN 60
DB 22 AKLVCFYFTNWAQYRQGEARFLPKDLPSCSLTHLIYAFAGMTNHQLSTTEWDETLYQSFN 81
QY 61 GLKXNPKLKTLLAIGGWNFGTQKFTDMVATANNRQTFVNSAIRFLKYSFDGLDWEY 120
DB 82 GLKXNPKLKTLLAIGGWNFGTQKFTDMVATANNRQTFVNSAIRFLKYSFDGLDWEY 141

QY 121 PGSQSPAVDKERFTTLVQDLANAFQEAQTSKGERLLLSAAVPAGQTVVDAGYEVDKIA 180
DB 142 PGSQSPAVDKERFTTLVQDLANAFQEAQTSKGERLLLSAAVPAGQTVVDAGYEVDKIA 201
QY 181 QNLDFVNLMAVDPHGWEKVTGHSNPLYKROESGAAASLNVDAAVQOQLQGTTPASKLI 240
DB 202 QNLDFVNLMAVDPHGWEKVTGHSNPLYKROESGAAASLNVDAAVQOQLQGTTPASKLI 261
QY 241 LGMPYGRSFTLASSSDTRVGAPATGSGTPGPTKEGGMALAYVEVCSWGKATKQRIQDOK 300
DB 262 LGMPYGRSFTLASSSDTRVGAPATGSGTPGPTKEGGMALAYVEVCSWGKATKQRIQDOK 321
QY 301 VPIYFRDNQWVGFDVSEFKTKVSYLKQKGLGAMVWALDLDLDFAGFSCNQGRYPLIQT 360
DB 322 VPIYFRDNQWVGFDVSEFKTKVSYLKQKGLGAMVWALDLDLDFAGFSCNQGRYPLIQT 381
QY 361 ROELSPLYPSPGTPPELVKPKQPSPEHGPSPGQDTFCQKADGLYPNPRRSSFYSKA 420
DB 382 ROELSPLYPSPGTPPELVKPKQPSPEHGPSPGQDTFCQKADGLYPNPRRSSFYSKA 441
QY 421 AGRLFQOQSCPTGLVFSNCKCCTWN 445
DB 442 AGRLFQOQSCPTGLVFSNCKCCTWN 466

RESULT 10

AAW31498
ID AAW31498 standard; Protein; 466 AA.

AC AAW31498;

XX 27-APR-1998 (first entry)

DE Human chitotriosidase variant.

XX Chitotriosidase; tissue remodelling disorder; diagnosis; therapy;
KW rheumatoid arthritis; atherosclerosis; human.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 102

FT /note= "encoded by RGC"

PN WO9736917-A1.

XX 09-OCT-1997.

XX 21-MAR-1997; 97WO-US05072.

XX 29-MAR-1996; 96US-0014295.

XX (HUMA-) HUMAN GENOME SCI INC.
PA (SMIK) SMITHKLINE BEECHAM CORP.

XX Chopra A, Gentz R, Kirkpatrick RB, Ni J, Thotakura NR;

XX WPI; 1997-503041/46.

DR N-PSDB; AAT89181.

XX New isolated human chitotriosidase gene - used to develop products
PT for the diagnosis and treatment of tissue remodeling disorders, e.g.
PT rheumatoid arthritis

XX Disclosure; Page 22-24; 34pp; English.

XX This protein comprises a variant of human chitotriosidase (see
CC also AAW31497). Its amino acid sequence was deduced from the
CC coding sequence of a full-length cDNA clone (see AAT89181). The
CC following are claimed: (1) a nucleic acid sequence encoding
CC chitotriosidase; (2) a method of diagnosing a tissue remodelling
CC disorder related to expression of a mutated chitotriosidase protein
CC in a host comprising carrying out nucleic acid amplification; and

CC (3) a method of detecting altered expression of a chitotriosidase
CC protein in a host comprising contacting a bodily sample with an
CC antibody. Chitotriosidase can degrade extracellular matrix
CC substrates with a similar carbohydrate structure to chitin. It can
CC be used to develop products which can be used in the diagnosis and
CC treatment of tissue remodelling disorders such as rheumatoid
CC arthritis or atherosclerosis.

XX SQ Sequence 466 AA;

Query Match 99.7%; Score 2391; DB 18; Length 466;

Best Local Similarity 99.8%; Pred. No. 3.6e-220;

Matches 444; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AKLVCFYNWAOYRQGEARFLPKDLDPSSLCTHLIIYAFAGMTNHQLSTTEWNETLIYQEFN 60

DB 22 AKLVCFYNWAOYRQGEARFLPKDLDPSSLCTHLIIYAFAGMTNHQLSTTEWNETLIYQEFN 81

QY 61 GLKKNPKLKTLLAIGGNWFGTKFTDMVATANNRQTFVNSAIRFLRYSPDGLDWEY 120

DB 82 GLKKNPKLKTLLAIGGNWFGTKFTDMVATANNRQTFVNSAIRFLRYSPDGLDWEY 141

QY 121 PGSQSPAVDKERFTTLVQDLANAFQEAQTSKGERLLLSAAVPAGQTVVDAGYEVDKIA 180

DB 142 PGSQSPAVDKERFTTLVQDLANAFQEAQTSKGERLLLSAAVPAGQTVVDAGYEVDKIA 201

QY 181 QNLDFVNLMAVDPHGWEKVTGHSNPLYKROESGAAASLNVDAAVQOQLQGTTPASKLI 240

DB 202 QNLDFVNLMAVDPHGWEKVTGHSNPLYKROESGAAASLNVDAAVQOQLQGTTPASKLI 261

QY 241 LGMPYGRSFTLASSSDTRVGAPATGSGTPGPTKEGGMALAYVEVCSWGKATKQRIQDOK 300

DB 262 LGMPYGRSFTLASSSDTRVGAPATGSGTPGPTKEGGMALAYVEVCSWGKATKQRIQDOK 321

QY 301 VPIYFRDNQWVGFDVSEFKTKVSYLKQKGLGAMVWALDLDLDFAGFSCNQGRYPLIQT 360

DB 322 VPIYFRDNQWVGFDVSEFKTKVSYLKQKGLGAMVWALDLDLDFAGFSCNQGRYPLIQT 381

QY 361 ROELSPLYPSPGTPPELVKPKQPSPEHGPSPGQDTFCQKADGLYPNPRRSSFYSKA 420

DB 382 ROELSPLYPSPGTPPELVKPKQPSPEHGPSPGQDTFCQKADGLYPNPRRSSFYSKA 441

QY 421 AGRLFQOQSCPTGLVFSNCKCCTWN 445

DB 442 AGRLFQOQSCPTGLVFSNCKCCTWN 466

RESULT 11

AAW31497

ID AAW31497 standard; Protein; 464 AA.

XX AAW31497;

XX 27-APR-1998 (first entry)

XX Human chitotriosidase.

XX Chitotriosidase; tissue remodelling disorder; diagnosis; therapy;
KW rheumatoid arthritis; atherosclerosis; human.

OS Homo sapiens.

XX WO9736917-A1.

XX 09-OCT-1997.

XX 21-MAR-1997; 97WO-US05072.

XX 29-MAR-1996; 96US-0014295.

XX (HUMA-) HUMAN GENOME SCI INC.
PA (SMIK) SMITHKLINE BEECHAM CORP.

XX

PI Chopra A, Gentz R, Kirkpatrick RB, Ni J, Thotakura NR;
 XX WPI: 1997-503041/46.
 DR N-PSDB; AAT89180.
 XX New isolated human chitotriosidase gene - used to develop products
 PT for the diagnosis and treatment of tissue remodeling disorders, e.g.
 PT rheumatoid arthritis
 XX
 PS Disclosure; Page 19-21; 34pp; English.
 XX
 CC This protein comprises human chitotriosidase. The amino acid
 CC sequence was deduced from the coding sequence of a full-length cDNA
 CC clone (see AAT89180) deposited as ATCC 69953. A preferred variant
 CC chitotriosidase sequence is also provided (see AAW31498). The
 CC following are claimed: (1) a nucleic acid sequence encoding
 CC chitotriosidase; (2) a method of diagnosing a tissue remodeling
 CC disorder related to expression of a mutated chitotriosidase protein
 CC in a host comprising carrying out nucleic acid amplification; and
 CC (3) a method of detecting altered expression of a chitotriosidase
 CC protein in a host comprising contacting a bodily sample with an
 CC antibody. Chitotriosidase can degrade extracellular matrix
 CC substrates with a similar carbohydrate structure to chitin. It can
 CC be used to develop products which can be used in the diagnosis and
 CC treatment of tissue remodeling disorders such as rheumatoid
 CC arthritis or atherosclerosis.
 XX
 SQ Sequence 464 AA;

Query Match 98.9%; Score 2372; DB 18; Length 464;
 Best Local Similarity 99.3%; Pred. No. 2.4e-218;
 Matches 442; Conservative 0; Mismatches 1; Indels 2; Gaps 1;
 QY 1 AKLVCFYFNWAGYRQGEARFLPKDLPSCLTLLIYAFAGMTNHQSLSTTEWDETLYQEFN 60
 DB 22 AKLVCFYFNWAGYRQGEARFLPKDLPSCLTLLIYAFAGMTNHQSLSTTEWDETLYQEFN 81
 QY 61 GLKKNPKLKTLLAIGGNWFGTKFTDMVATANNRQTFVNSAIRFLRYKSPDGLDWEY 120
 DB 82 GLKKNPKLKTLLAIGGNWFGTKFTDMVATANNRQTFVNSAIRFLRYKSPDGLDWEY 141
 QY 121 PGSQSPAVDKERFTTLVQDLANAFQGEAQTSGKRLLSAAVPAGQTYVDAGYVDKIA 180
 DB 142 PGSQSPAVDKERFTTL--DLANAFQGEAQTSGKRLLSAAVPAGQTYVDAGYVDKIA 199
 QY 181 QNLDFVNLMAVDYFHGSWEKVTGHNSPLYKROESGAASLNVDAAVQWLOKGTTPASKLI 240
 DB 200 QNLDFVNLMAVDYFHGSWEKVTGHNSPLYKROESGAASLNVDAAVQWLOKGTTPASKLI 259
 QY 241 LGMPYGRSFTLLASSSDTRVGAPATGSGTPPGFTKEGGMALAYEVCWKGATKQRIQDOK 300
 DB 260 LGMPYGRSFTLLASSSDTRVGAPATGSGTPPGFTKEGGMALAYEVCWKGATKQRIQDOK 319
 QY 301 VPIYFRDNQWGFDDVESFKTKVSLYKQKGLGAMVWALDLDFFAGFSCNQGRIYPLIQT 360
 DB 320 VPIYFRDNQWGFDDVESFKTKVSLYKQKGLGAMVWALDLDFFAGFSCNQGRIYPLIQT 379
 QY 361 ROELSLPYLPSTGTPLEVPKGPQSEPEHGRSPGQDTFCOGKADGLYNPRRSSFYSKA 420
 DB 380 ROELSLPYLPSTGTPLEVPKGPQSEPEHGRSPGQDTFCOGKADGLYNPNRRSSFYSKA 439
 QY 421 AGRLFOQSCPTGLVFSNSCKCCTWN 445
 DB 440 AGRLFOQSCPTGLVFSNSCKCCTWN 464
 RESULT 12
 AAW40261
 ID AAW40261 standard; Protein; 373 AA.
 XX
 AC AAW40261;
 XX
 DT 15-JUN-1998 (first entry)

XX Human chitinase protein fragment.
 DE Chitinase; human; fungal infection; immunogen; diagnosis; treatment;
 XX Chitinase; disease; transgenic; detection; hybridisation; antifungal;
 KW Gaucher's disease; arthritis; overexpression; extracellular matrix.
 KW
 XX Homo sapiens.
 OS
 XX WO9747752-A1.
 PN
 XX 18-DEC-1997.
 PD
 XX 16-JUN-1997; 97WO-US10460.
 PF
 XX 14-JUN-1996; 96US-0663618.
 PR
 XX (ICOS-) ICOS CORP.
 PA
 XX Gray PW;
 PI
 XX WPI: 1998-052316/05.
 DR N-PSDB; AAV10435.
 DR
 XX Nucleic acids encoding human chitinase - useful as antifungal
 PT agents, especially in combination with other antifungals
 PT
 XX Claim 23; Page 48-49; 63pp; English.
 PS
 XX This sequence encodes a fragment of a novel human chitinase protein
 CC which lacks the C-terminal 72 residues of the mature protein. These
 CC residues are not critical to chitinase enzymatic activity. Chitinases are
 CC useful for treating or preventing fungal infection and as immunogens for
 CC generating antibodies which are used to purify, detect and quantify
 CC chitinases, e.g. for diagnosis of Gaucher's disease. The nucleic acid
 CC sequence of the chitinase is also useful as a probe to identify and
 CC isolate genomic DNA encoding chitinases or similar proteins, or cells
 CC expressing them or to generate transgenic ('knockout') rodents. It can
 CC also be used in hybridisation assays and to detect genetic alterations
 CC in the chitinase gene related to disease. Agents that inhibit this
 CC protein may be useful in treatment of Gaucher's disease and rheumatoid
 CC arthritis, where overexpression of the protein can damage the
 CC extracellular matrix. Chitinase also improves the activity of other
 CC antifungal agents and may allow a reduction in the dose of such agents,
 CC and thus of their side effects.
 XX
 SQ Sequence 373 AA;

Query Match 82.7%; Score 1982; DB 19; Length 373;
 Best Local Similarity 100.0%; Pred. No. 4.3e-181;
 Matches 373; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AKLVCFYFNWAGYRQGEARFLPKDLPSCLTLLIYAFAGMTNHQSLSTTEWDETLYQEFN 60
 DB 1 AKLVCFYFNWAGYRQGEARFLPKDLPSCLTLLIYAFAGMTNHQSLSTTEWDETLYQEFN 60
 QY 61 GLKKNPKLKTLLAIGGNWFGTKFTDMVATANNRQTFVNSAIRFLRYKSPDGLDWEY 120
 DB 61 GLKKNPKLKTLLAIGGNWFGTKFTDMVATANNRQTFVNSAIRFLRYKSPDGLDWEY 120
 QY 121 PGSQSPAVDKERFTTLVQDLANAFQGEAQTSGKRLLSAAVPAGQTYVDAGYVDKIA 180
 DB 121 PGSQSPAVDKERFTTLVQDLANAFQGEAQTSGKRLLSAAVPAGQTYVDAGYVDKIA 180
 QY 181 QNLDFVNLMAVDYFHGSWEKVTGHNSPLYKROESGAASLNVDAAVQWLOKGTTPASKLI 240
 DB 181 QNLDFVNLMAVDYFHGSWEKVTGHNSPLYKROESGAASLNVDAAVQWLOKGTTPASKLI 240
 QY 241 LGMPYGRSFTLLASSSDTRVGAPATGSGTPPGFTKEGGMALAYEVCWKGATKQRIQDOK 300
 DB 241 LGMPYGRSFTLLASSSDTRVGAPATGSGTPPGFTKEGGMALAYEVCWKGATKQRIQDOK 300
 QY 301 VPIYFRDNQWGFDDVESFKTKVSLYKQKGLGAMVWALDLDFFAGFSCNQGRIYPLIQT 360

Db 301 VPYIFRDQWVGFDVESFKTKVSYLKQKGLGAMVWALDLDFAFGSCNQGRYPLIQTL 360
 QY 361 ROELSPLPLPSGT 373
 Db 361 ROELSPLPLPSGT 373
 RESULT 13
 AAY42427
 ID AAY42427 standard; Protein; 373 AA.
 AC AAY42427;
 DT 10-DEC-1999 (first entry)
 XX
 DE Clone of the C-terminal fragment of human chitinase.
 XX
 KW chitin; fungal infection; immunocompromised; AIDS; chemotherapy;
 KW organ transplant; parasite; chitin-binding; allele; vector;
 KW truncated protein; bacterial expression.
 XX
 OS Synthetic.
 OS Homo sapiens.
 PN WO9946390-A1.
 XX
 PD 16-SEP-1999.
 XX
 PF 12-MAR-1999; 99WO-US05343.
 XX
 PR 12-MAR-1998; 98US-0039198.
 XX
 PA (ICOS-) ICOS CORP.
 XX
 PI Gray PW, Tjoelker LW;
 XX WPI; 1999-551417/46.
 DR
 XX Novel chitin-binding fragments of human chitinase used to treat fungal
 PT infections in animals -
 XX
 PS Example 5; Page 66-67; 83pp; English.
 XX
 CC This is the amino acid sequence of the C-terminal fragment of human
 CC chitinase, which can be expressed in bacterial cells. The fragment was
 CC prepared by amplifying the MO-218 plasmid with primers 218-1 (AAZ21855)
 CC and T-END (AAZ21856).
 CC Chitinase fragments can be used to screen for proteins or other
 CC molecules that specifically bind to the chitin-binding domain of human
 CC chitinase or that modulate its activity. These compounds are useful for
 CC immunization, as well as for purifying chitinase, as well as for
 CC detection and quantification of chitinase. Polynucleotide fragments of
 CC the invention are useful as a source of probes and primers, and to
 CC express the proteins recombinantly. The chitinase fragments, when
 CC conjugated to antifungal compounds, are used to treat animals,
 CC especially humans, infected with chitin-containing parasites such as
 CC fungi. Fungal infection treated include candidiasis, aspergillosis,
 CC coccidioidomycosis, blastomycosis, paracoccidioidomycosis,
 CC mucormycosis, histoplasmosis, cryptococcosis, chromoblastomycosis,
 CC sporotrichosis, and dermatophytoses.
 CC Chitin can be degraded by the enzyme chitinase. Use of whole chitinase
 CC protein for treating infections, especially fungal infections, is
 CC problematic. In view of the increasing incidents of life-threatening
 CC fungal infection in e.g. immunocompromised individuals, there exists a
 CC need for identifying new compounds for treating fungal infection. The
 CC chitin-binding fragments of the present invention provide this need.
 XX
 SQ Sequence 373 AA;
 Query Match 82.7%; Score 1982; DB 20; Length 373;
 Best Local Similarity 100.0%; Pred. No. 4.3e-181;
 Matches 373; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKLVCFYFNWAGYRQGEARFLPKDLDPSTLCTHLIYAFAGMTNHQLSTTEWNETLYQEFN 60
 Db 1 AKLVCFYFNWAGYRQGEARFLPKDLDPSTLCTHLIYAFAGMTNHQLSTTEWNETLYQEFN 60
 QY 61 GLKKNPKLKTLLAIGGNWFGTKFTDMVATANNRQTFFNSAIRFLRKYSPDGLDLDWEY 120
 Db 61 GLKKNPKLKTLLAIGGNWFGTKFTDMVATANNRQTFFNSAIRFLRKYSPDGLDLDWEY 120
 QY 121 PGSQSPAVDKERFTTLVODLANAFQOEATQSGKERLLLSAAVPAQOTVVDAGYEVDKIA 180
 Db 121 PGSQSPAVDKERFTTLVODLANAFQOEATQSGKERLLLSAAVPAQOTVVDAGYEVDKIA 180
 QY 181 QNLDFVNLMAFYDFHGSWEKVTCHNSPLYKROESGAASLNVDAAVQOQLKGTTPASKLI 240
 Db 181 QNLDFVNLMAFYDFHGSWEKVTCHNSPLYKROESGAASLNVDAAVQOQLKGTTPASKLI 240
 QY 241 LGMPTVGRSFTLASSSDTRVGAPATGSGTPGFTKEGGMLAYEYVCSWKGATKORIQQOK 300
 Db 241 LGMPTVGRSFTLASSSDTRVGAPATGSGTPGFTKEGGMLAYEYVCSWKGATKORIQQOK 300
 QY 301 VPIYIFRDQWVGFDVESFKTKVSYLKQKGLGAMVWALDLDFAFGSCNQGRYPLIQTL 360
 Db 301 VPIYIFRDQWVGFDVESFKTKVSYLKQKGLGAMVWALDLDFAFGSCNQGRYPLIQTL 360
 QY 361 ROELSPLPLPSGT 373
 Db 361 ROELSPLPLPSGT 373
 RESULT 14
 ABB76293
 ID ABB76293 standard; Protein; 373 AA.
 XX ABB76293;
 AC ABB76293;
 DT 12-AUG-2002 (first entry)
 XX
 DE Human chitinase truncated polypeptide.
 XX
 KW Chitinase; enzyme; human; fungicide; antifungal; infection;
 KW candidiasis; aspergillosis; coccidioidomycosis; blastomycosis;
 KW paracoccidioidomycosis; histoplasmosis; cryptococcosis;
 KW chromoblastomycosis; sporotrichosis; mucormycosis; dermatophytosis;
 KW Pneumocystis.
 XX
 OS Homo sapiens.
 XX
 PN US6372212-B1.
 XX
 PD 16-APR-2002.
 XX
 PF 16-JUN-1997; 97US-0877599.
 XX
 PR 14-JUN-1996; 96US-0663618.
 XX
 PA (ICOS-) ICOS CORP.
 XX
 PI Gray PW;
 XX WPI; 2002-442449/47.
 DR
 XX Co-administering chitinase to improve the effectiveness of fungicidal
 PT drugs e.g. amphotericin B or itraconazole, useful for treating fungal
 PT infections e.g. candidiasis, coccidioidomycosis and blastomycosis -
 XX
 XX Example 5; Column 49-42; 26pp; English.
 XX
 CC The present sequence is the protein sequence of a human mature
 CC chitinase truncated polypeptide, in which the 72 C-terminal amino
 CC acids of the native mature protein (see ABB76291) are deleted.
 CC The coding region for the polypeptide was obtained by PCR from
 CC chitinase MO-218 cDNA (see ABL57380), and the polypeptide was

CC produced in transfected COS cells. This 39 kDa polypeptide lacks
CC 6 cysteine residues of the native protein yet retains similar
CC specific enzymatic activity. The invention provides human
CC chitinase polynucleotides and polypeptides, and methods for the
CC recombinant production of human chitinase products, which are
CC expected to be useful for treating fungal infections. The
CC chitinase can be used to improve the activity of a non-chitinase
CC antifungal agent in the treatment of candidiasis, aspergillosis,
CC coccidioidomycosis, blastomycosis, paracoccidioidomycosis,
CC histoplasmosis, cryptococcosis, chromoblastomycosis, sporotrichosis,
CC mucormycosis, dermatophytoses and Pneumocystis infections.
XX
SQ Sequence 373 AA;

Query Match 82.7%; Score 1982; DB 23; Length 373;
Best Local Similarity 100.0%; Pred. No. 4.3e-181;
Matches 373; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AKLVCFYTNWAQYRQGEARFLPKDLDPSCSLCTHLYAFAGMTHQSLSTTEWNETLYQEFN 60
DB 1 AKLVCFYTNWAQYRQGEARFLPKDLDPSCSLCTHLYAFAGMTHQSLSTTEWNETLYQEFN 60
QY 61 GLKKNPKLKTLLAIGGWNFGTKFTDMVATANNRQTFVNSAIRFLRKYSPDGLDLDWEY 120
DB 61 GLKKNPKLKTLLAIGGWNFGTKFTDMVATANNRQTFVNSAIRFLRKYSPDGLDLDWEY 120
QY 121 PGSQSPAVDKERFTTLVQDLANAFQEAQTSKGERLLLSAAVPAGQTYVDAGYVDKIA 180
DB 121 PGSQSPAVDKERFTTLVQDLANAFQEAQTSKGERLLLSAAVPAGQTYVDAGYVDKIA 180
QY 181 QNLDFVNLMAVDPHGSEKVTGHNPSLYKRBESGAAASLVNDAVQWLGKGTDPASKLI 240
DB 181 QNLDFVNLMAVDPHGSEKVTGHNPSLYKRBESGAAASLVNDAVQWLGKGTDPASKLI 240
QY 241 LGMPTYGRSFTLASSDTRVGAPATGSGTPGPTKEGGMLAYEYVCSWGATKQRIQDOK 300
DB 241 LGMPTYGRSFTLASSDTRVGAPATGSGTPGPTKEGGMLAYEYVCSWGATKQRIQDOK 300
QY 301 VPIYIFRDNQWGFDDVESFKTKVSYLKQKGLGGMVWALDLDFFAGFSCNQGRIYPIQTL 360
DB 301 VPIYIFRDNQWGFDDVESFKTKVSYLKQKGLGGMVWALDLDFFAGFSCNQGRIYPIQTL 360
QY 361 ROELSLPYLPST 373
DB 361 ROELSLPYLPST 373

RESULT 15
AAW40262
ID AAW40262 standard; Protein; 373 AA.
XX
AC AAW40262;
XX
DE Human chitinase protein analogue.
XX
KW Chitinase; human; fungal infection; immunogen; diagnosis; treatment;
KW Gaucher's disease; transgenic; detection; hybridisation; antifungal;
KW Rheumatoid arthritis; overexpression; extracellular matrix.
XX
OS Homo sapiens.
OS Synthetic.
FH Key Location/Qualifiers
FT Misc-difference 370 /label= P370S
FT /note= "Wild type Pro is replaced by Ser"
XX
PN W09747752-A1.
XX
PD 18-DEC-1997.

PF 16-JUN-1997; 97WO-US10460.
XX
PR 14-JUN-1996; 96US-0663618.
XX
PA (ICOS-) ICOS CORP.
XX
PI Gray PW;
XX
DR WPI; 1998-052316/05.
XX
PT Nucleic acids encoding human chitinase - useful as antifungal
XX agents, especially in combination with other antifungals
PS Claim 29; Page 49-50; 63pp; English.
XX
CC This sequence encodes a fragment of a novel human chitinase protein
CC analogue in which a proline at position 370 of the wild type protein
CC (see AAW40261) is substituted with a serine and the C-terminal 72
CC residues of the mature protein are deleted. These residues are not
CC critical to chitinase enzymatic activity. Chitinases are useful for
CC treating or preventing fungal infection and as immunogens for generating
CC antibodies which are used to purify, detect and quantify chitinases, e.g.
CC for diagnosis of Gaucher's disease. The nucleic acid sequence of the
CC chitinase is also useful as a probe to identify and isolate genomic DNA
CC encoding chitinases or similar proteins, or cells expressing them or to
CC generate transgenic ('knockout') rodents. It can also be used in
CC hybridisation assays and to detect genetic alterations in the chitinase
CC gene related to disease. Agents that inhibit this protein may be useful
CC in treatment of Gaucher's disease and rheumatoid arthritis, where
CC overexpression of the protein can damage the extracellular matrix.
CC Chitinase also improves the activity of other antifungal agents and may
CC allow a reduction in the dose of such agents, and thus of their side
CC effects.
XX
SQ Sequence 373 AA;

Query Match 82.3%; Score 1974; DB 19; Length 373;
Best Local Similarity 99.7%; Pred. No. 2.5e-180;
Matches 372; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 AKLVCFYTNWAQYRQGEARFLPKDLDPSCSLCTHLYAFAGMTHQSLSTTEWNETLYQEFN 60
DB 1 AKLVCFYTNWAQYRQGEARFLPKDLDPSCSLCTHLYAFAGMTHQSLSTTEWNETLYQEFN 60
QY 61 GLKKNPKLKTLLAIGGWNFGTKFTDMVATANNRQTFVNSAIRFLRKYSPDGLDLDWEY 120
DB 61 GLKKNPKLKTLLAIGGWNFGTKFTDMVATANNRQTFVNSAIRFLRKYSPDGLDLDWEY 120
QY 121 PGSQSPAVDKERFTTLVQDLANAFQEAQTSKGERLLLSAAVPAGQTYVDAGYVDKIA 180
DB 121 PGSQSPAVDKERFTTLVQDLANAFQEAQTSKGERLLLSAAVPAGQTYVDAGYVDKIA 180
QY 181 QNLDFVNLMAVDPHGSEKVTGHNPSLYKRBESGAAASLVNDAVQWLGKGTDPASKLI 240
DB 181 QNLDFVNLMAVDPHGSEKVTGHNPSLYKRBESGAAASLVNDAVQWLGKGTDPASKLI 240
QY 241 LGMPTYGRSFTLASSDTRVGAPATGSGTPGPTKEGGMLAYEYVCSWGATKQRIQDOK 300
DB 241 LGMPTYGRSFTLASSDTRVGAPATGSGTPGPTKEGGMLAYEYVCSWGATKQRIQDOK 300
QY 301 VPIYIFRDNQWGFDDVESFKTKVSYLKQKGLGGMVWALDLDFFAGFSCNQGRIYPIQTL 360
DB 301 VPIYIFRDNQWGFDDVESFKTKVSYLKQKGLGGMVWALDLDFFAGFSCNQGRIYPIQTL 360
QY 361 ROELSLPYLPST 373
DB 361 ROELSLPYLPST 373

Search completed: June 29, 2003, 20:59:40
Job time : 30.6228 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2003, 20:57:59 ; Search time 11.52 Seconds
(without alignments)
1136.565 Million cell updates/sec

Title: US-10-004-219b-10
Perfect score: 2398
Sequence: 1 AKLVCFYFTWVAQYRQGEARF.....QQSCPTGLVFSNCKCCTWN 445

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/prodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/prodata/1/iaa/6CTUS_COMB.pep:*
6: /cgn2_6/prodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2398	100.0	466	2	US-08-486-839-4
2	2398	100.0	466	3	US-09-151-011-4
3	2398	100.0	466	4	US-09-039-198A-2
4	2398	100.0	466	4	US-09-343-623-4
5	2398	100.0	466	4	US-08-877-599-2
6	2398	100.0	466	4	US-09-267-574-2
7	2392	99.7	466	4	US-09-039-198A-4
8	2392	99.7	466	4	US-08-877-599-4
9	2392	99.7	466	4	US-09-267-574-4
10	1982	82.7	373	4	US-09-039-198A-14
11	1982	82.7	373	4	US-08-877-599-14
12	1982	82.7	373	4	US-09-267-574-14
13	1974	82.3	373	4	US-09-039-198A-15
14	1974	82.3	373	4	US-08-877-599-15
15	1974	82.3	373	4	US-09-267-574-15
16	1935	80.7	387	2	US-08-486-839-6
17	1935	80.7	387	3	US-09-151-011-6
18	1935	80.7	387	4	US-09-343-623-6
19	1045.5	43.6	385	2	US-08-694-915-2
20	1045.5	43.6	416	2	US-08-694-915-4
21	1021.5	42.6	383	4	US-09-459-749D-17
22	824	34.4	554	3	US-08-524-051-2
23	803.5	34.4	554	3	US-09-052-778-16
24	803.5	33.5	559	4	US-09-545-814-14
25	803.5	33.5	583	4	US-09-545-814-2
26	803.5	33.5	583	4	US-09-545-814-5
27	797.5	33.3	635	4	US-09-545-814-32

28	762	31.8	490	4	US-09-292-225-41	Sequence 41, Appl
29	762	31.8	509	4	US-09-292-225-35	Sequence 35, Appl
30	762	31.8	509	4	US-09-292-225-38	Sequence 38, Appl
31	738	30.8	536	4	US-09-292-225-21	Sequence 21, Appl
32	738	30.8	555	4	US-09-292-225-15	Sequence 15, Appl
33	738	30.8	555	4	US-09-292-225-18	Sequence 18, Appl
34	660	27.5	489	4	US-09-545-814-29	Sequence 29, Appl
35	499.5	20.8	389	1	US-07-939-501A-1	Sequence 1, Appl
36	499.5	20.8	389	4	US-08-448-398-7	Sequence 7, Appl
37	499.5	20.8	423	1	US-07-939-501A-10	Sequence 10, Appl
38	499.5	20.8	423	1	US-07-939-501A-12	Sequence 12, Appl
39	486	20.3	424	1	US-08-045-269C-2	Sequence 2, Appl
40	486	20.3	424	3	US-08-371-680-2	Sequence 2, Appl
41	485	20.2	424	5	PCT-US94-01198-2	Sequence 2, Appl
42	474	19.8	442	3	US-09-052-778-2	Sequence 2, Appl
43	462.5	19.3	377	2	US-08-591-629-8	Sequence 8, Appl
44	460.5	19.2	371	2	US-08-591-629-2	Sequence 2, Appl
45	444.5	18.5	440	3	US-09-052-778-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1
US-08-486-839-4
; Sequence 4, Application US/08486839
; Patent No. 5928928
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: A human chitinase, its recombinant
; TITLE OF INVENTION: production, its use for decomposing chitin, its use
; TITLE OF INVENTION: in therapy or prophylaxis against infection diseases.
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11758
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,839
; FILING DATE: 07 - June - 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Baron, Ronald J.
; REGISTRATION NUMBER: 29,281
; REFERENCE/DOCKET NUMBER: 294-26
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 466 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; US-08-486-839-4

Query Match 100.0%; Score 2398; DB 2; Length 466;
Best Local Similarity 100.0%; Pred. No. 5,1e-218;
Matches 445; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKLVCFYFTWVAQYRQGEARFLPKDLPSCLTSLIYAFAGTWHQSTWENDETLQEFN 60
|||
DB 22 AKLVCFYFTWVAQYRQGEARFLPKDLPSCLTSLIYAFAGTWHQSTWENDETLQEFN 81

Query Match 100.0%; Score 2398; DB 3; Length 466;
Best Local Similarity 100.0%; Pred. No. 5,1e-218;
Matches 445; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 GLKKNPKLTKLLAIGGNFTGQKFTDMVATANNRQTFVNSAIRFLRKYSFGDGLDLIDWEY 120
DB 82 GLKKNPKLTKLLAIGGNFTGQKFTDMVATANNRQTFVNSAIRFLRKYSFGDGLDLIDWEY 141

QY 121 PGSQSPAVDKERFTTLVQDLANAFQOEATSGKERLLLSAAVPAGQTYVDAGYVDKIA 180
DB 142 PGSQSPAVDKERFTTLVQDLANAFQOEATSGKERLLLSAAVPAGQTYVDAGYVDKIA 201

QY 181 QNLDFVNLMAIDFHGSWEKVTGHNPSPLYKQESGAAASLNVDAAVQOVLQKGTTPASKLI 240
DB 202 QNLDFVNLMAIDFHGSWEKVTGHNPSPLYKQESGAAASLNVDAAVQOVLQKGTTPASKLI 261

QY 241 LGMPTYGRSFTLASSSDTRVGAPATGSGTPGPTKEGGMLAYEYVCSWKGATKQRIQDOK 300
DB 262 LGMPTYGRSFTLASSSDTRVGAPATGSGTPGPTKEGGMLAYEYVCSWKGATKQRIQDOK 321

QY 301 VPIYFRDNQWGFDDVESFKTKVSYLKQKGLGAMVWALDLDFFAGFSCNQGRIPLIQTLL 360
DB 322 VPIYFRDNQWGFDDVESFKTKVSYLKQKGLGAMVWALDLDFFAGFSCNQGRIPLIQTLL 381

QY 361 ROELSPLYPSTGTPLEVPKQPSPEHGPSPGQDTFCQKADGLYPNPRSSSFYSCA 420
DB 382 ROELSPLYPSTGTPLEVPKQPSPEHGPSPGQDTFCQKADGLYPNPRSSSFYSCA 441

QY 421 AGRLFQOQSCPTGLVFSNSCKCCTWN 445
DB 442 AGRLFQOQSCPTGLVFSNSCKCCTWN 466

RESULT 2

US-09-151-011-4
; Sequence 4, Application US/09151011
; Patent No. 6057142
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: A Human Chitinase, Its Recombinant
; TITLE OF INVENTION: Production, Its Use For Decomposing Chitin, Its Use in
; TITLE OF INVENTION: Therapy or Prophylaxis Against Infection Diseases.
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Hoffmann & Baron, LLP
; STREET: 6900 Jericho Turnpike
; CITY: Syosset
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11791
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; SOFTWARE:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/151,011
; FILING DATE: 10 - September - 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Morris, Robert C.
; REGISTRATION NUMBER: 42,910
; REFERENCE/DOCKET NUMBER: 294-32 DIV
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 466 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; US-09-151-011-4

RESULT 3

US-09-039-198A-2
; Sequence 2, Application US/09039198A
; Patent No. 6200951
; GENERAL INFORMATION:
; APPLICANT: Gray, Patrick W.
; APPLICANT: Tjoelker, Larry W.
; TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive/6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/039,198A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 27866/34391
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 466 amino acids
; TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-039-198A-2

Query Match 100.0%; Score 2398; DB 4; Length 466;
Best Local Similarity 100.0%; Pred. No. 5.1e-218;
Matches 445; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 AKLVCFYFTNWAQYRQGEARFLPKDLDPSCSLTHLIYAFAGMTNHQSLSTTEWDETLYQEFN 60
22 AKLVCFYFTNWAQYRQGEARFLPKDLDPSCSLTHLIYAFAGMTNHQSLSTTEWDETLYQEFN 81
61 GLKKNPKLKTLLAIGGNFQKFTDMVATANNRQTFVNSAIRFLRYSPDGLDWEY 120
82 GLKKNPKLKTLLAIGGNFQKFTDMVATANNRQTFVNSAIRFLRYSPDGLDWEY 141
121 PGSQSPAVDKERFTLLVQDLANAFQQAOTSGKERLLLSAAVPAGQTYVDAGYEVDKIA 180
142 PGSQSPAVDKERFTLLVQDLANAFQQAOTSGKERLLLSAAVPAGQTYVDAGYEVDKIA 201
181 QNLDFVNLMAVDPHGSEWEKVTGHNSPLYKROESGAAASLNVDAVQWLOKGTTPASKLI 240
202 QNLDFVNLMAVDPHGSEWEKVTGHNSPLYKROESGAAASLNVDAVQWLOKGTTPASKLI 261
241 LGMPTYGRSFTLASSSDTRVGAPATGSGTGPFTKEGGMLAYEYVCSWKATKQRIQDOK 300
262 LGMPTYGRSFTLASSSDTRVGAPATGSGTGPFTKEGGMLAYEYVCSWKATKQRIQDOK 321
301 VPIYFRDQWGVDFDVESEFKTKVSYLKOKGLGGAMVWALDDDFAGFCNQGRIYPLIOTL 360
322 VPIYFRDQWGVDFDVESEFKTKVSYLKOKGLGGAMVWALDDDFAGFCNQGRIYPLIOTL 381
361 ROELSLPLPSCGTPELEVPKQPSPEHGPSPGQDTFCQKADGLYNPNRERSFYSCA 420
382 ROELSLPLPSCGTPELEVPKQPSPEHGPSPGQDTFCQKADGLYNPNRERSFYSCA 441
421 AGRLFQSCPTGLVFSNCKCCTWN 445
442 AGRLFQSCPTGLVFSNCKCCTWN 466

RESULT 4
US-09-343-623-4
Sequence 4, Application US/09343623
Patent No. 6303118
GENERAL INFORMATION:
APPLICANT: A human chitinase, its recombinant
TITLE OF INVENTION: production, its use for decomposing chitin, its use
TITLE OF INVENTION: in therapy or prophylaxis against infection diseases.
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hoffmann & Baron
STREET: 350 Jericho Turnpike
CITY: Jericho
STATE: New York
COUNTRY: United States of America
ZIP: 11758
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/343,623
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/486,839
FILING DATE: 07-June-1995
ATTORNEY/AGENT INFORMATION:
NAME: Baron, Ronald J.
REGISTRATION NUMBER: 29,281

REFERENCE/DOCKET NUMBER: 294-26
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 822-3550
TELEFAX: (516) 822-3582
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 466 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-09-343-623-4

Query Match 100.0%; Score 2398; DB 4; Length 466;
Best Local Similarity 100.0%; Pred. No. 5.1e-218;
Matches 445; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 AKLVCFYFTNWAQYRQGEARFLPKDLDPSCSLTHLIYAFAGMTNHQSLSTTEWDETLYQEFN 60
22 AKLVCFYFTNWAQYRQGEARFLPKDLDPSCSLTHLIYAFAGMTNHQSLSTTEWDETLYQEFN 81
61 GLKKNPKLKTLLAIGGNFQKFTDMVATANNRQTFVNSAIRFLRYSPDGLDWEY 120
82 GLKKNPKLKTLLAIGGNFQKFTDMVATANNRQTFVNSAIRFLRYSPDGLDWEY 141
121 PGSQSPAVDKERFTLLVQDLANAFQQAOTSGKERLLLSAAVPAGQTYVDAGYEVDKIA 180
142 PGSQSPAVDKERFTLLVQDLANAFQQAOTSGKERLLLSAAVPAGQTYVDAGYEVDKIA 201
181 QNLDFVNLMAVDPHGSEWEKVTGHNSPLYKROESGAAASLNVDAVQWLOKGTTPASKLI 240
202 QNLDFVNLMAVDPHGSEWEKVTGHNSPLYKROESGAAASLNVDAVQWLOKGTTPASKLI 261
241 LGMPTYGRSFTLASSSDTRVGAPATGSGTGPFTKEGGMLAYEYVCSWKATKQRIQDOK 300
262 LGMPTYGRSFTLASSSDTRVGAPATGSGTGPFTKEGGMLAYEYVCSWKATKQRIQDOK 321
301 VPIYFRDQWGVDFDVESEFKTKVSYLKOKGLGGAMVWALDDDFAGFCNQGRIYPLIOTL 360
322 VPIYFRDQWGVDFDVESEFKTKVSYLKOKGLGGAMVWALDDDFAGFCNQGRIYPLIOTL 381
361 ROELSLPLPSCGTPELEVPKQPSPEHGPSPGQDTFCQKADGLYNPNRERSFYSCA 420
382 ROELSLPLPSCGTPELEVPKQPSPEHGPSPGQDTFCQKADGLYNPNRERSFYSCA 441
421 AGRLFQSCPTGLVFSNCKCCTWN 445
442 AGRLFQSCPTGLVFSNCKCCTWN 466

RESULT 5
US-08-877-599-2
Sequence 2, Application US/08877599
Patent No. 6372212
GENERAL INFORMATION:
APPLICANT: Gray, Patrick W.
TITLE OF INVENTION: Chitinase Materials and Methods
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/877,599

```
/ FILING DATE:
/ CLASSIFICATION: 514
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/663,618
/ FILING DATE: 14-JUN-1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Rin-Laures, Li-Hsien
/ REGISTRATION NUMBER: 33,547
/ REFERENCE/DOCKET NUMBER: 27866/33994
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 312/474-6300
/ TELEFAX: 312/474-0448
/ TELEX: 25-3856
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 466 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-877-599-2

Query Match          100.0%; Score 2398; DB 4; Length 466;
Best Local Similarity 100.0%; Pred. No. 5.1e-218;
Matches 445; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKLVCFYFTNWAQYRQGEARFLPKDLPDLPSLCTHLIYAFAGTNHQLSTTEWDETLYQEFN 60
DB 22 AKLVCFYFTNWAQYRQGEARFLPKDLPDLPSLCTHLIYAFAGTNHQLSTTEWDETLYQEFN 81
QY 61 GLKKNPKLKTLLAIGGWNFGTKFTDMVATANNROTFFVNSAIRFLRKYSFDCGLDLDWEY 120
DB 82 GLKKNPKLKTLLAIGGWNFGTKFTDMVATANNROTFFVNSAIRFLRKYSFDCGLDLDWEY 141
QY 121 PGSQSPAVDKERFTTLVQDLANAFQGEAQTSGKERLLLSAAVPAGQTYVDAGYEVDKIA 180
DB 142 PGSQSPAVDKERFTTLVQDLANAFQGEAQTSGKERLLLSAAVPAGQTYVDAGYEVDKIA 201
QY 181 QNLDFVNLMAFYDPHGSWEKVTGHSNPLYKROESGAAASLNVDAAVQOVLKGTTPASKLI 240
DB 202 QNLDFVNLMAFYDPHGSWEKVTGHSNPLYKROESGAAASLNVDAAVQOVLKGTTPASKLI 261
QY 241 LGMPTYGRSFTLLASSSDTRVGAPATGSGTPGPTKEGGMLAYEYVCSWKGATKQRIQDOK 300
DB 262 LGMPTYGRSFTLLASSSDTRVGAPATGSGTPGPTKEGGMLAYEYVCSWKGATKQRIQDOK 321
QY 301 VPIYFRDNQWVGDDVESFKTKVSYLKQKGLGAMVWALDLDLDDFAGFSCNQGRIPIQTL 360
DB 322 VPIYFRDNQWVGDDVESFKTKVSYLKQKGLGAMVWALDLDLDDFAGFSCNQGRIPIQTL 381
QY 361 ROELSPLYPSPGTPPELVKPGQSPPEHGPSPGQDTFCQKADGLYPNPERSSFYSYCA 420
DB 382 ROELSPLYPSPGTPPELVKPGQSPPEHGPSPGQDTFCQKADGLYPNPERSSFYSYCA 441
QY 421 AGRLFQOCSPTGLVFSNSCKCTWN 445
DB 442 AGRLFQOCSPTGLVFSNSCKCTWN 466

RESULT 6
US-09-267-574-2
/ Sequence 2, Application US/09267574
/ Patent No. 6399571
/ GENERAL INFORMATION:
/ APPLICANT: Gray, Patrick W.
/ TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
/ FILE REFERENCE: 27866/35407
/ CURRENT APPLICATION NUMBER: US/09/267,574
/ EARLIER FILING DATE: 1999-03-12
/ EARLIER APPLICATION NUMBER: 09/039,198
/ NUMBER OF SEQ ID NOS: 39
/ SOFTWARE: Patentin Ver. 2.0
```

```
/ SEQ ID NO 2
/ LENGTH: 466
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ US-09-267-574-2

Query Match          100.0%; Score 2398; DB 4; Length 466;
Best Local Similarity 100.0%; Pred. No. 5.1e-218;
Matches 445; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKLVCFYFTNWAQYRQGEARFLPKDLPDLPSLCTHLIYAFAGTNHQLSTTEWDETLYQEFN 60
DB 22 AKLVCFYFTNWAQYRQGEARFLPKDLPDLPSLCTHLIYAFAGTNHQLSTTEWDETLYQEFN 81
QY 61 GLKKNPKLKTLLAIGGWNFGTKFTDMVATANNROTFFVNSAIRFLRKYSFDCGLDLDWEY 120
DB 82 GLKKNPKLKTLLAIGGWNFGTKFTDMVATANNROTFFVNSAIRFLRKYSFDCGLDLDWEY 141
QY 121 PGSQSPAVDKERFTTLVQDLANAFQGEAQTSGKERLLLSAAVPAGQTYVDAGYEVDKIA 180
DB 142 PGSQSPAVDKERFTTLVQDLANAFQGEAQTSGKERLLLSAAVPAGQTYVDAGYEVDKIA 201
QY 181 QNLDFVNLMAFYDPHGSWEKVTGHSNPLYKROESGAAASLNVDAAVQOVLKGTTPASKLI 240
DB 202 QNLDFVNLMAFYDPHGSWEKVTGHSNPLYKROESGAAASLNVDAAVQOVLKGTTPASKLI 261
QY 241 LGMPTYGRSFTLLASSSDTRVGAPATGSGTPGPTKEGGMLAYEYVCSWKGATKQRIQDOK 300
DB 262 LGMPTYGRSFTLLASSSDTRVGAPATGSGTPGPTKEGGMLAYEYVCSWKGATKQRIQDOK 321
QY 301 VPIYFRDNQWVGDDVESFKTKVSYLKQKGLGAMVWALDLDLDDFAGFSCNQGRIPIQTL 360
DB 322 VPIYFRDNQWVGDDVESFKTKVSYLKQKGLGAMVWALDLDLDDFAGFSCNQGRIPIQTL 381
QY 361 ROELSPLYPSPGTPPELVKPGQSPPEHGPSPGQDTFCQKADGLYPNPERSSFYSYCA 420
DB 382 ROELSPLYPSPGTPPELVKPGQSPPEHGPSPGQDTFCQKADGLYPNPERSSFYSYCA 441
QY 421 AGRLFQOCSPTGLVFSNSCKCTWN 445
DB 442 AGRLFQOCSPTGLVFSNSCKCTWN 466

RESULT 7
US-09-039-198A-4
/ Sequence 4, Application US/09039198A
/ Patent No. 6200951
/ GENERAL INFORMATION:
/ APPLICANT: Gray, Patrick W.
/ APPLICANT: Tjoelker, Larry W.
/ TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
/ NUMBER OF SEQUENCES: 34
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
/ STREET: 233 South Wacker Drive/6300 Sears Tower
/ CITY: Chicago
/ STATE: Illinois
/ COUNTRY: United States of America
/ ZIP: 60606-6402
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/039,198A
/ FILING DATE:
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Rin-Laures, Li-Hsien
/ REGISTRATION NUMBER: 33,547
/ REFERENCE/DOCKET NUMBER: 27866/34391
/ TELECOMMUNICATION INFORMATION:
```

TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 466 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-039-198A-4

Query Match 99.7%; Score 2392; DB 4; Length 466;
Best Local Similarity 99.8%; Pred. No. 1.9e-217;
Matches 444; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AKLVCFYFTNWAQYRQGEARFLPKDLPDLPSLCTHLIIYAFAGMTNHQSLSTTEWNETLYQEFN 60
DB 22 AKLVCFYFTNWAQYRQGEARFLPKDLPDLPSLCTHLIIYAFAGMTNHQSLSTTEWNETLYQEFN 81
QY 61 GLKKNPKLKTLLAIGGWNFGTKFTDMVATANNRQTFVNSAIRFLRKYSFGLDLWDEY 120
DB 82 GLKKNPKLKTLLAIGGWNFGTKFTDMVATANNRQTFVNSAIRFLRKYSFGLDLWDEY 141
QY 121 PGSQSPAVDKERFTLVODLANAQOEATSGKERLLLSAAVPAQOTYVDAGYEVDKIA 180
DB 142 PGSQSPAVDKERFTLVODLANAQOEATSGKERLLLSAAVPAQOTYVDAGYEVDKIA 201
QY 181 QNLDFVNLMAVDFHGSWEKVTGHNSPLYKROESGAAASLNVDAVQWLOKGTTPASKLI 240
DB 202 QNLDFVNLMAVDFHGSWEKVTGHNSPLYKROESGAAASLNVDAVQWLOKGTTPASKLI 261
QY 241 LGMPTYGRSFTLLASSSDTRVGAPATGSGTPGFTKEGGMALAYEVCWKGATKQRIQDOK 300
DB 262 LGMPTYGRSFTLLASSSDTRVGAPATGSGTPGFTKEGGMALAYEVCWKGATKQRIQDOK 321
QY 301 VPIYFRDNQWGFDDVESFKTKVYLKQKGLGAMVWALDDDFAGFSCNQGRYPLIOTL 360
DB 322 VPIYFRDNQWGFDDVESFKTKVYLKQKGLGAMVWALDDDFAGFSCNQGRYPLIOTL 381
QY 361 ROELSLPYLPSTGTPLEVPKQPESEPHGSPGQDTFCQKADGLYNPNRPRSSFYSCA 420
DB 442 AGRLFOQSCPTGLVFSNSCKCTWN 466

RESULT 8

US-08-877-599-4

Sequence 4, Application US/08877599

Patent No. 6372212

GENERAL INFORMATION:

APPLICANT: Gray, Patrick W.

TITLE OF INVENTION: Chitinase Materials and Methods

NUMBER OF SEQUENCES: 17

CORRESPONDENCE ADDRESS:

ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun

STREET: 6300 Sears Tower, 233 South Wacker Drive

CITY: Chicago

STATE: Illinois

COUNTRY: United States of America

ZIP: 60606-6402

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/877,599

FILING DATE:

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/663,618

FILING DATE: 14-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Rin-Laures, Li-Hsien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 27866/33994
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 466 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-877-599-4

Query Match 99.7%; Score 2392; DB 4; Length 466;
Best Local Similarity 99.8%; Pred. No. 1.9e-217;
Matches 444; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AKLVCFYFTNWAQYRQGEARFLPKDLPDLPSLCTHLIIYAFAGMTNHQSLSTTEWNETLYQEFN 60
DB 22 AKLVCFYFTNWAQYRQGEARFLPKDLPDLPSLCTHLIIYAFAGMTNHQSLSTTEWNETLYQEFN 81
QY 61 GLKKNPKLKTLLAIGGWNFGTKFTDMVATANNRQTFVNSAIRFLRKYSFGLDLWDEY 120
DB 82 GLKKNPKLKTLLAIGGWNFGTKFTDMVATANNRQTFVNSAIRFLRKYSFGLDLWDEY 141
QY 121 PGSQSPAVDKERFTLVODLANAQOEATSGKERLLLSAAVPAQOTYVDAGYEVDKIA 180
DB 142 PGSQSPAVDKERFTLVODLANAQOEATSGKERLLLSAAVPAQOTYVDAGYEVDKIA 201
QY 181 QNLDFVNLMAVDFHGSWEKVTGHNSPLYKROESGAAASLNVDAVQWLOKGTTPASKLI 240
DB 202 QNLDFVNLMAVDFHGSWEKVTGHNSPLYKROESGAAASLNVDAVQWLOKGTTPASKLI 261
QY 241 LGMPTYGRSFTLLASSSDTRVGAPATGSGTPGFTKEGGMALAYEVCWKGATKQRIQDOK 300
DB 262 LGMPTYGRSFTLLASSSDTRVGAPATGSGTPGFTKEGGMALAYEVCWKGATKQRIQDOK 321
QY 301 VPIYFRDNQWGFDDVESFKTKVYLKQKGLGAMVWALDDDFAGFSCNQGRYPLIOTL 360
DB 322 VPIYFRDNQWGFDDVESFKTKVYLKQKGLGAMVWALDDDFAGFSCNQGRYPLIOTL 381
QY 361 ROELSLPYLPSTGTPLEVPKQPESEPHGSPGQDTFCQKADGLYNPNRPRSSFYSCA 420
DB 382 ROELSLPYLPSTGTPLEVPKQPESEPHGSPGQDTFCQKADGLYNPNRPRSSFYSCA 441
QY 421 AGRLFOQSCPTGLVFSNSCKCTWN 445
DB 442 AGRLFOQSCPTGLVFSNSCKCTWN 466

RESULT 9

US-09-267-574-4

Sequence 4, Application US/09267574

Patent No. 6399571

GENERAL INFORMATION:

APPLICANT: Gray, Patrick W.

APPLICANT: Tjoelker, Larry W.

TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS

FILE REFERENCE: 27866/35407

CURRENT APPLICATION NUMBER: US/09/267,574

CURRENT FILING DATE: 1999-03-12

EARLIER APPLICATION NUMBER: 09/039,198

EARLIER FILING DATE: 1998-03-12

NUMBER OF SEQ ID NOS: 39

SOFTWARE: Patent In Ver. 2.0

SEQ ID NO 4

LENGTH: 466

TYPE: PRT

ORGANISM: Homo sapiens

US-09-267-574-4

Query Match 99.7%; Score 2392; DB 4; Length 466;
Best Local Similarity 99.8%; Pred. No. 1.9e-217;
Matches 444; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AKLVCFYFTNWAQYRQGEARFLPKDLPSCSLTHLIYAFAGMTNHQSLSTTEWNETLYOEFN 60
DB 22 AKLVCFYFTNWAQYRQGEARFLPKDLPSCSLTHLIYAFAGMTNHQSLSTTEWNETLYOEFN 81
QY 61 GLKKNPKLKTLLAIGGNWFGTKFTDMVATANNRQTFVNSAIRFLRYKVSFGLDLDWEY 120
DB 82 GLKKNPKLKTLLAIGGNWFGTKFTDMVATANNRQTFVNSAIRFLRYKVSFGLDLDWEY 141
QY 121 PGSGGSPAVDKERFTTLVQDLANAFQEAQTSKERLLLSAAVPAGQTYVDAGYEVDKIA 180
DB 142 PGSGGSPAVDKERFTTLVQDLANAFQEAQTSKERLLLSAAVPAGQTYVDAGYEVDKIA 201
QY 181 QNLDFVNLMAVDYFHGSWEKVTGHNPSPLYKQESGAAASLNVDAAVQOQLKGTTPASKLI 240
DB 202 QNLDFVNLMAVDYFHGSWEKVTGHNPSPLYKQESGAAASLNVDAAVQOQLKGTTPASKLI 261
QY 241 LGMPTVGRSFTLASSSDTRVGAPATSGTPGPTKEGGLMAYEVCWSKGATKQRIQDOK 300
DB 262 LGMPTVGRSFTLASSSDTRVGAPATSGTPGPTKEGGLMAYEVCWSKGATKQRIQDOK 321
QY 301 VPIYFRDNQWGFDDVESFKTKVSYLKQKGLGAMVWALDLDLDDFAGFSCNQGRIYPLIOTL 360
DB 322 VPIYFRDNQWGFDDVESFKTKVSYLKQKGLGAMVWALDLDLDDFAGFSCNQGRIYPLIOTL 381
QY 361 ROELSPLYPSPGTPELEVPKPGQSPBEPHGPSQDTPFCQKADGLYPNPRLSSFYSCA 420
DB 382 ROELSPLYPSPGTPELEVPKPGQSPBEPHGPSQDTPFCQKADGLYPNPRLSSFYSCA 441
QY 421 AGRLFQSCPTGLVFNSSCKCCTWN 445
DB 442 AGRLFQSCPTGLVFNSSCKCCTWN 466

RESULT 10

US-09-039-198A-14
; Sequence 14, Application US/09039198A
; Patent No. 6200951
; GENERAL INFORMATION:
; APPLICANT: Gray, Patrick W.
; APPLICANT: Toelker, Larry W.
; TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive/6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/039,198A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 27866/34391
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:

LENGTH: 373 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLSCULE TYPE: peptide
US-09-039-198A-14

Query Match 82.7%; Score 1982; DB 4; Length 373;

Best Local Similarity 100.0%; Pred. No. 7.1e-179;
Matches 373; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKLVCFYFTNWAQYRQGEARFLPKDLPSCSLTHLIYAFAGMTNHQSLSTTEWNETLYOEFN 60
DB 1 AKLVCFYFTNWAQYRQGEARFLPKDLPSCSLTHLIYAFAGMTNHQSLSTTEWNETLYOEFN 60
QY 61 GLKKNPKLKTLLAIGGNWFGTKFTDMVATANNRQTFVNSAIRFLRYKVSFGLDLDWEY 120
DB 61 GLKKNPKLKTLLAIGGNWFGTKFTDMVATANNRQTFVNSAIRFLRYKVSFGLDLDWEY 120
QY 121 PGSGGSPAVDKERFTTLVQDLANAFQEAQTSKERLLLSAAVPAGQTYVDAGYEVDKIA 180
DB 121 PGSGGSPAVDKERFTTLVQDLANAFQEAQTSKERLLLSAAVPAGQTYVDAGYEVDKIA 180
QY 181 QNLDFVNLMAVDYFHGSWEKVTGHNPSPLYKQESGAAASLNVDAAVQOQLKGTTPASKLI 240
DB 181 QNLDFVNLMAVDYFHGSWEKVTGHNPSPLYKQESGAAASLNVDAAVQOQLKGTTPASKLI 240
QY 241 LGMPTVGRSFTLASSSDTRVGAPATSGTPGPTKEGGLMAYEVCWSKGATKQRIQDOK 300
DB 241 LGMPTVGRSFTLASSSDTRVGAPATSGTPGPTKEGGLMAYEVCWSKGATKQRIQDOK 300
QY 301 VPIYFRDNQWGFDDVESFKTKVSYLKQKGLGAMVWALDLDLDDFAGFSCNQGRIYPLIOTL 360
DB 301 VPIYFRDNQWGFDDVESFKTKVSYLKQKGLGAMVWALDLDLDDFAGFSCNQGRIYPLIOTL 360
QY 361 ROELSPLYPSPGT 373
DB 361 ROELSPLYPSPGT 373

RESULT 11

US-08-877-599-14
; Sequence 14, Application US/08877599
; Patent No. 6372212
; GENERAL INFORMATION:
; APPLICANT: Gray, Patrick W.
; TITLE OF INVENTION: Chitinase Materials and Methods
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/877,599
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/663,618
; FILING DATE: 14-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 27866/33994
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300

TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 373 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-877-599-14

Query Match 82.7%; Score 1982; DB 4; Length 373;
Best Local Similarity 100.0%; Pred. No. 7.1e-179;
Matches 373; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKLVCFYFTNWAQYRQGEARFLPKDLPDLPSLCTHLIYAFAGMTNHQLSTTEWNETIYQEFN 60
DB 1 AKLVCFYFTNWAQYRQGEARFLPKDLPDLPSLCTHLIYAFAGMTNHQLSTTEWNETIYQEFN 60
QY 61 GLKXNPKLKTLLAIGGNFQKFTDMVATANNRQTFVNSAIRFLRKYSPDGLDWEY 120
DB 61 GLKXNPKLKTLLAIGGNFQKFTDMVATANNRQTFVNSAIRFLRKYSPDGLDWEY 120
QY 121 PGSQSPAVDKERFTTLVODLANAQOEAOQTSKGERLLLSAAVPAGQTYVDAGYEVDKIA 180
DB 121 PGSQSPAVDKERFTTLVODLANAQOEAOQTSKGERLLLSAAVPAGQTYVDAGYEVDKIA 180
QY 181 QNLDFVNLMAVDYFHGSWEKVTGHNSPLYKROESGAASLNVDAAVQWLOKGTTPASKLI 240
DB 181 QNLDFVNLMAVDYFHGSWEKVTGHNSPLYKROESGAASLNVDAAVQWLOKGTTPASKLI 240
QY 241 LGMPTYGRSFTLLASSDTRVGAPATGSGTPGPFKEGGMLAYEYVCSWKGATKQRIQDOK 300
DB 241 LGMPTYGRSFTLLASSDTRVGAPATGSGTPGPFKEGGMLAYEYVCSWKGATKQRIQDOK 300
QY 301 VPIYFRDNQWGFDDVESFKTVSVLKQKGLGAMVWALDLDLDDFAGFSCNOGRYPLIOTL 360
DB 301 VPIYFRDNQWGFDDVESFKTVSVLKQKGLGAMVWALDLDLDDFAGFSCNOGRYPLIOTL 360
QY 361 ROELSLPYLPSTGT 373
DB 361 ROELSLPYLPSTGT 373

RESULT 12
US-09-267-574-14
Sequence 14, Application US/09267574
Patent No. 6399571
GENERAL INFORMATION:

APPLICANT: Gray, Patrick W.
TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
FILE REFERENCE: 27866/35407
CURRENT APPLICATION NUMBER: US/09/267,574
CURRENT FILING DATE: 1999-03-12
EARLIER APPLICATION NUMBER: 09/039,198
EARLIER FILING DATE: 1998-03-12
NUMBER OF SEQ ID NOS: 39
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 14
LENGTH: 373
TYPE: PRT
ORGANISM: Homo sapiens
US-09-267-574-14

Query Match 82.7%; Score 1982; DB 4; Length 373;
Best Local Similarity 100.0%; Pred. No. 7.1e-179;
Matches 373; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKLVCFYFTNWAQYRQGEARFLPKDLPDLPSLCTHLIYAFAGMTNHQLSTTEWNETIYQEFN 60
DB 1 AKLVCFYFTNWAQYRQGEARFLPKDLPDLPSLCTHLIYAFAGMTNHQLSTTEWNETIYQEFN 60

QY 61 GLKXNPKLKTLLAIGGNFQKFTDMVATANNRQTFVNSAIRFLRKYSPDGLDWEY 120
DB 61 GLKXNPKLKTLLAIGGNFQKFTDMVATANNRQTFVNSAIRFLRKYSPDGLDWEY 120
QY 121 PGSQSPAVDKERFTTLVODLANAQOEAOQTSKGERLLLSAAVPAGQTYVDAGYEVDKIA 180
DB 121 PGSQSPAVDKERFTTLVODLANAQOEAOQTSKGERLLLSAAVPAGQTYVDAGYEVDKIA 180
QY 181 QNLDFVNLMAVDYFHGSWEKVTGHNSPLYKROESGAASLNVDAAVQWLOKGTTPASKLI 240
DB 181 QNLDFVNLMAVDYFHGSWEKVTGHNSPLYKROESGAASLNVDAAVQWLOKGTTPASKLI 240
QY 241 LGMPTYGRSFTLLASSDTRVGAPATGSGTPGPFKEGGMLAYEYVCSWKGATKQRIQDOK 300
DB 241 LGMPTYGRSFTLLASSDTRVGAPATGSGTPGPFKEGGMLAYEYVCSWKGATKQRIQDOK 300
QY 301 VPIYFRDNQWGFDDVESFKTVSVLKQKGLGAMVWALDLDLDDFAGFSCNOGRYPLIOTL 360
DB 301 VPIYFRDNQWGFDDVESFKTVSVLKQKGLGAMVWALDLDLDDFAGFSCNOGRYPLIOTL 360
QY 361 ROELSLPYLPSTGT 373
DB 361 ROELSLPYLPSTGT 373

RESULT 13

US-09-039-198A-15
Sequence 15, Application US/09039198A
Patent No. 6200951
GENERAL INFORMATION:

APPLICANT: Gray, Patrick W.
APPLICANT: Tjoelker, Larry W.
TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive/6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/039,198A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Rin-Laures, Li-Hsien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 27866/34391
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 373 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-039-198A-15

Query Match 82.3%; Score 1974; DB 4; Length 373;
Best Local Similarity 99.7%; Pred. No. 4e-178;
Matches 372; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AKLVCFYFTNWAQYRQGEARFLPKDLPDLPSLCTHLIYAFAGMTNHQLSTTEWNETIYQEFN 60
DB 1 AKLVCFYFTNWAQYRQGEARFLPKDLPDLPSLCTHLIYAFAGMTNHQLSTTEWNETIYQEFN 60

QY 61 GLKKNPKLTLAIGGNFQKFTDMVATANNRQTFVNSAIRFLRKYSFGLDLDWEY 120
DB 61 GLKKNPKLTLAIGGNFQKFTDMVATANNRQTFVNSAIRFLRKYSFGLDLDWEY 120
QY 121 PGSQSPAVDKERFTTLVODLANAFOEAOQTSKERLLLSAAVPAGQTYVDAGYVDKIA 180
DB 121 PGSQSPAVDKERFTTLVODLANAFOEAOQTSKERLLLSAAVPAGQTYVDAGYVDKIA 180
QY 181 QNLDFVNLMAIDFHGSWEKVTGHSNPLYKROBESGAAASLNVDAAVQOQLKGTTPASKLI 240
DB 181 QNLDFVNLMAIDFHGSWEKVTGHSNPLYKROBESGAAASLNVDAAVQOQLKGTTPASKLI 240
QY 241 LGMPTTGRSFTLASSSDTRVGPATGSGTGPFTKEGGLMAYEVCWGKATKQRIQDOK 300
DB 241 LGMPTTGRSFTLASSSDTRVGPATGSGTGPFTKEGGLMAYEVCWGKATKQRIQDOK 300
QY 301 VPIFRDNOWGFDVESFRTKVSYLKQGLGAMWALDLDFFAGFSCNQGRYPLIOTL 360
DB 301 VPIFRDNOWGFDVESFRTKVSYLKQGLGAMWALDLDFFAGFSCNQGRYPLIOTL 360
QY 361 ROELSPLYPSSGT 373
DB 361 ROELSPLYPSSGT 373

RESULT 14

US-08-877-599-15
; Sequence 15, Application US/08877599
; Patent No. 6372212
; GENERAL INFORMATION:
; APPLICANT: Gray, Patrick W.
; TITLE OF INVENTION: Chitinase Materials and Methods
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/877,599
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/663,618
; FILING DATE: 14-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 27866/33994
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 373 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-877-599-15

Query Match 82.3%; Score 1974; DB 4; Length 373;
Best Local Similarity 99.7%; Pred. No. 4e-178;
Matches 372; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AKLVCFYTNWAQYRQGEARFLPKDLPDLPSLCTHLIYAFAGMTNHQLSTTEWNETLYQEFN 60
DB 1 AKLVCFYTNWAQYRQGEARFLPKDLPDLPSLCTHLIYAFAGMTNHQLSTTEWNETLYQEFN 60
QY 61 GLKKNPKLTLAIGGNFQKFTDMVATANNRQTFVNSAIRFLRKYSFGLDLDWEY 120
DB 61 GLKKNPKLTLAIGGNFQKFTDMVATANNRQTFVNSAIRFLRKYSFGLDLDWEY 120
QY 121 PGSQSPAVDKERFTTLVODLANAFOEAOQTSKERLLLSAAVPAGQTYVDAGYVDKIA 180
DB 121 PGSQSPAVDKERFTTLVODLANAFOEAOQTSKERLLLSAAVPAGQTYVDAGYVDKIA 180
QY 181 QNLDFVNLMAIDFHGSWEKVTGHSNPLYKROBESGAAASLNVDAAVQOQLKGTTPASKLI 240
DB 181 QNLDFVNLMAIDFHGSWEKVTGHSNPLYKROBESGAAASLNVDAAVQOQLKGTTPASKLI 240
QY 241 LGMPTTGRSFTLASSSDTRVGPATGSGTGPFTKEGGLMAYEVCWGKATKQRIQDOK 300
DB 241 LGMPTTGRSFTLASSSDTRVGPATGSGTGPFTKEGGLMAYEVCWGKATKQRIQDOK 300
QY 301 VPIFRDNOWGFDVESFRTKVSYLKQGLGAMWALDLDFFAGFSCNQGRYPLIOTL 360
DB 301 VPIFRDNOWGFDVESFRTKVSYLKQGLGAMWALDLDFFAGFSCNQGRYPLIOTL 360
QY 361 ROELSPLYPSSGT 373
DB 361 ROELSPLYPSSGT 373

RESULT 15

US-09-267-574-15
; Sequence 15, Application US/09267574
; Patent No. 6399571
; GENERAL INFORMATION:
; APPLICANT: Gray, Patrick W.
; TITLE OF INVENTION: CHITINASE
; FILE REFERENCE: 27866/35407
; CURRENT APPLICATION NUMBER: US/09/267,574
; EARLIER FILING DATE: 1999-03-12
; EARLIER APPLICATION NUMBER: 09/039,198
; EARLIER FILING DATE: 1998-03-12
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 15
; LENGTH: 373
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-267-574-15

Query Match 82.3%; Score 1974; DB 4; Length 373;

Best Local Similarity 99.7%; Pred. No. 4e-178;
Matches 372; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AKLVCFYTNWAQYRQGEARFLPKDLPDLPSLCTHLIYAFAGMTNHQLSTTEWNETLYQEFN 60
DB 1 AKLVCFYTNWAQYRQGEARFLPKDLPDLPSLCTHLIYAFAGMTNHQLSTTEWNETLYQEFN 60
QY 61 GLKKNPKLTLAIGGNFQKFTDMVATANNRQTFVNSAIRFLRKYSFGLDLDWEY 120
DB 61 GLKKNPKLTLAIGGNFQKFTDMVATANNRQTFVNSAIRFLRKYSFGLDLDWEY 120
QY 121 PGSQSPAVDKERFTTLVODLANAFOEAOQTSKERLLLSAAVPAGQTYVDAGYVDKIA 180
DB 121 PGSQSPAVDKERFTTLVODLANAFOEAOQTSKERLLLSAAVPAGQTYVDAGYVDKIA 180
QY 181 QNLDFVNLMAIDFHGSWEKVTGHSNPLYKROBESGAAASLNVDAAVQOQLKGTTPASKLI 240
DB 181 QNLDFVNLMAIDFHGSWEKVTGHSNPLYKROBESGAAASLNVDAAVQOQLKGTTPASKLI 240
QY 241 LGMPTTGRSFTLASSSDTRVGPATGSGTGPFTKEGGLMAYEVCWGKATKQRIQDOK 300

Db 241 LCMPTYGRSFTLASSSDTRVGAPATGSGTPGPFTKEGGMLAYYEVCSWKGATKQRIQDOK 300
Qy 301 VFYIFRDNQVGFDDVESEFKTKVSYLKQKGLGAMWALDLDLDDFAGFSCNQGRYPLIOTL 360
Db 301 VFYIFRDNQVGFDDVESEFKTKVSYLKQKGLGAMWALDLDLDDFAGFSCNQGRYPLIOTL 360
Qy 361 ROELSLPYLPSGT 373
Db 361 ROELSLPYLSSGT 373

Search completed: June 29, 2003, 21:03:29
Job time : 12.52 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2003, 21:01:56 ; Search time 17.1154 Seconds
(without alignments)
2850.818 Million cell updates/sec

Title: US-10-004-219B-10
Perfect score: 2398
Sequence: 1 AKLVCFYTNWAGYRGEARF.....QQSCPTGLVFNCKCTWN 445

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 424699 seqs, 109646833 residues

Total number of hits satisfying chosen parameters: 424699

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US03_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2398	100.0	445	9	US-10-004-219B-10
2	1323	55.2	455	9	US-10-004-219B-10
3	1323	55.2	476	9	US-10-004-219B-10
4	1301.5	54.3	452	9	US-10-004-219B-9
5	1301.5	54.3	473	9	US-10-004-219B-4
6	1041.5	43.4	383	9	US-10-097-340-45
7	1021.5	42.6	383	10	US-09-459-749D-17
8	762	31.8	490	9	US-10-218-743-41
9	762	31.8	509	9	US-10-218-743-35
10	762	31.8	509	9	US-10-218-743-38
11	738	30.8	536	9	US-10-218-743-21
12	738	30.8	555	9	US-10-218-743-15
13	738	30.8	555	9	US-10-218-743-18
14	487	20.3	371	9	US-09-923-844B-2
15	335.5	14.0	376	10	US-09-748-033-3
16	214	8.9	170	9	US-10-218-743-44
17	101.5	4.2	1289	10	US-09-738-363-4
18	101	4.2	411	9	US-09-866-050A-302
19	100	4.2	802	10	US-09-757-049A-1

20	97.5	4.1	346	9	US-10-260-877-80	Sequence 80, Appl
21	97	4.0	393	9	US-09-977-418-12	Sequence 12, Appl
22	97	4.0	393	9	US-09-977-033A-12	Sequence 12, Appl
23	97	4.0	393	10	US-09-840-787-11	Sequence 11, Appl
24	96.5	4.0	1276	10	US-09-982-610-24	Sequence 24, Appl
25	96	4.0	416	9	US-10-023-282-461	Sequence 461, Appl
26	94	3.9	289	9	US-09-957-187-14	Sequence 14, Appl
27	94	3.9	1154	9	US-10-224-624-8	Sequence 8, Appl
28	92.5	3.9	298	10	US-09-748-033-1	Sequence 1, Appl
29	92	3.8	363	9	US-10-023-282-278	Sequence 278, Appl
30	91.5	3.8	1140	9	US-09-974-973-19	Sequence 19, Appl
31	91.5	3.8	1140	9	US-09-738-626-4265	Sequence 4265, Appl
32	91.5	3.8	1140	9	US-10-045-072-2	Sequence 2, Appl
33	91.5	3.8	1157	9	US-09-974-973-2	Sequence 4, Appl
34	91.5	3.8	1157	9	US-09-974-973-4	Sequence 4, Appl
35	91.5	3.8	1190	9	US-09-964-295-2	Sequence 2, Appl
36	91.5	3.8	2322	9	US-09-919-039-15	Sequence 15, Appl
37	91	3.8	408	9	US-10-027-806-74	Sequence 74, Appl
38	91	3.8	408	9	US-10-034-623-74	Sequence 74, Appl
39	91	3.8	408	9	US-10-027-801-74	Sequence 74, Appl
40	91	3.8	1293	10	US-09-815-242-10079	Sequence 10079, A
41	90.5	3.8	2783	10	US-09-816-669A-14	Sequence 14, Appl
42	90	3.8	2971	9	US-10-146-473-50	Sequence 50, Appl
43	89.5	3.7	222	9	US-09-727-855B-5	Sequence 5, Appl
44	89.5	3.7	1079	9	US-10-112-488-39	Sequence 39, Appl
45	89	3.7	235	9	US-09-738-626-5245	Sequence 5245, Appl

ALIGNMENTS

RESULT 1
US-10-004-219B-10
; Sequence 10, Application US/10004219B
; Publication No. US20030087414A1
; GENERAL INFORMATION:
; APPLICANT: Macrozyme
; APPLICANT: Aerts, Johannes M.F.G.
; APPLICANT: Boot, Rolf G.
; TITLE OF INVENTION: A mammalian mucinase, its recombinant production, and its use in therapy or prophylaxis against diseases in which mucus is involved or infection diseases
; FILE REFERENCE: 2183-51360S
; CURRENT APPLICATION NUMBER: US/10/004,219B
; CURRENT FILING DATE: 2001-11-02
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 10
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1)..(445)
; OTHER INFORMATION: /note="Human chitotriosidase"
US-10-004-219B-10

Query Match	100.0%;	Score 2398;	DB 9;	Length 445;
Best Local Similarity	100.0%;	Pred. No. 5.2e-202;	Mismatches 0;	Indels 0;
Matches 445;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	AKLVCFYTNWAGYRGEARFLPKDLPDLCTHLIYAFAGMTNHLSTTEWNETLYQEFN	60	
DB	1	AKLVCFYTNWAGYRGEARFLPKDLPDLCTHLIYAFAGMTNHLSTTEWNETLYQEFN	60	
QY	61	GLKXNPKLTLATGGWNFGTKETDMVATANNRQTFVNSAIRFLRKYSPGLDLDWEY	120	
DB	61	GLKXNPKLTLATGGWNFGTKETDMVATANNRQTFVNSAIRFLRKYSPGLDLDWEY	120	
QY	121	PGSQSPAVDKERFTTLVODLANAQOEAOQTSQKRLLSAAVPAGQTVYDAGYVDKIA	180	
DB	121	PGSQSPAVDKERFTTLVODLANAQOEAOQTSQKRLLSAAVPAGQTVYDAGYVDKIA	180	

181 QNLDVNLMAIDYDFHGSWEKVTGHSNPLYKQESGAAASLNVDAAVOQMLQKGTTPASKLI 240
181 QNLDVNLMAIDYDFHGSWEKVTGHSNPLYKQESGAAASLNVDAAVOQMLQKGTTPASKLI 240
241 LGMPTTGRSFTLASSSDTRVGA PATGSGTGPPTKGGMLAYEVCWSKGATKQRIODQK 300
241 LGMPTTGRSFTLASSSDTRVGA PATGSGTGPPTKGGMLAYEVCWSKGATKQRIODQK 300
301 VPIYFRDQWVGDDVESFKTKVSYLKQKGLGGAMWALDLDLDPAGFSCNQGRIPIOTL 360
301 VPIYFRDQWVGDDVESFKTKVSYLKQKGLGGAMWALDLDLDPAGFSCNQGRIPIOTL 360
361 ROELSPLPSPGTPPELEVPKGPSEPEHGPSPGQDTFCQKADGLYPNPRSSFYSCA 420
361 ROELSPLPSPGTPPELEVPKGPSEPEHGPSPGQDTFCQKADGLYPNPRSSFYSCA 420
421 AGRLFQOCSCTGLVFNNSCKCCTWN 445
421 AGRLFQOCSCTGLVFNNSCKCCTWN 445

RESULT 2
US-10-004-219B-14
; Sequence 14, Application US/10004219B
; Publication No. US20030087414A1
; GENERAL INFORMATION:
; APPLICANT: Macrozyme
; APPLICANT: Aerts, Johannes M.F.G.
; APPLICANT: Boot, Rolf G.
; TITLE OF INVENTION: A mammalian mucinase, its recombinant production, and
; TITLE OF INVENTION: its use in therapy or prophylaxis against diseases in
; TITLE OF INVENTION: which mucus is involved or infection diseases
; FILE REFERENCE: 2183-5136US
; CURRENT APPLICATION NUMBER: US/10/004,219B
; CURRENT FILING DATE: 2001-11-02
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1)..(455)
; OTHER INFORMATION: /note="Human AMCase"
US-10-004-219B-14

Query Match 55.2%; Score 1323; DB 9; Length 455;
Best Local Similarity 52.2%; Pred. No. 1.2e-107;
Matches 240; Conservative 70; Mismatches 126; Indels 24; Gaps 3;

QY 2 KLVCFYTNWAQYRQGEARFLPKDLPCLTHLIYAFAGTNHQLSTTEWDETLVQEFNG 61
DB 2 QLTCTYTNWAQYRQGEARFLPKDLPCLTHLIYAFAGTNHQLSTTEWDETLVQEFNG 61
QY 62 LKKNPKLTLAIGGWNFGTQKFTDMVATANNROTFFVNSAIRFLRKYSFGDGLDWEYP 121
DB 62 LKKNPKLTLAIGGWNFGTQKFTDMVATANNROTFFVNSAIRFLRKYSFGDGLDWEYP 121
QY 122 GSQGSFPAVDKERFTTLVQDLANAFQEAQTSKGERLLLSAAVPAQGTYYDAGYEVDKIAQ 181
DB 122 GSQGSFPAVDKERFTTLVQDLANAFQEAQTSKGERLLLSAAVPAQGTYYDAGYEVDKIAQ 181
QY 182 NLDVNLMAIDYDFHGSWEKVTGHSNPLYKQESGAAASLNVDAAVOQMLQKGTTPASKLI 241
DB 182 NLDVNLMAIDYDFHGSWEKVTGHSNPLYKQESGAAASLNVDAAVOQMLQKGTTPASKLI 241
QY 242 GMPTTGRSFTLASSSDTRVGA PATGSGTGPPTKGGMLAYEVCWS--KGATKQRIODQ 299
DB 242 GMPTTGRSFTLASSSDTRVGA PATGSGTGPPTKGGMLAYEVCWS--KGATKQRIODQ 299
QY 300 KVPYIFRDQWVGDDVESFKTKVSYLKQKGLGGAMWALDLDLDPAGFSCNQGRIPIOT 359
DB 300 KVPYIFRDQWVGDDVESFKTKVSYLKQKGLGGAMWALDLDLDPAGFSCNQGRIPIOT 359

Db 302 EVPYAYQGNVWVGVDNIKSFDIKAQWLKHNFKGAMVWALDLDLDTGTTCNOCKFPLIST 361
QY 360 LRQELSLPVLPSGTPELEVPKGPQSEP-----EHGPPSGQDTFCQKAD 404
Db 362 LRKAL-----GLOSASCTAPAQPIEPITAAPSGSGSGSGSGSGSGGFCFAVRAN 414
QY 405 GLYPNPRSSFYSCAAGRLFOQSCPTGLVFNNSCKCCTW 444
Db 415 GLYPVANNRNFHWCVNGVTTYQONCOAGLVFDTSCDCCNW 454

RESULT 3
US-10-004-219B-1
; Sequence 1, Application US/10004219B
; Publication No. US20030087414A1
; GENERAL INFORMATION:
; APPLICANT: Macrozyme
; APPLICANT: Aerts, Johannes M.F.G.
; APPLICANT: Boot, Rolf G.
; TITLE OF INVENTION: A mammalian mucinase, its recombinant production, and
; TITLE OF INVENTION: its use in therapy or prophylaxis against diseases in
; TITLE OF INVENTION: which mucus is involved or infection diseases
; FILE REFERENCE: 2183-5136US
; CURRENT APPLICATION NUMBER: US/10/004,219B
; CURRENT FILING DATE: 2001-11-02
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: human AMCase
; OTHER INFORMATION: amino acid sequence deduced from cDNA sequence
US-10-004-219B-1

Query Match 55.2%; Score 1323; DB 9; Length 476;
Best Local Similarity 52.2%; Pred. No. 1.2e-107;
Matches 240; Conservative 70; Mismatches 126; Indels 24; Gaps 3;

QY 2 KLVCFYTNWAQYRQGEARFLPKDLPCLTHLIYAFAGTNHQLSTTEWDETLVQEFNG 61
DB 23 QLTCTYTNWAQYRQGEARFLPKDLPCLTHLIYAFAGTNHQLSTTEWDETLVQEFNG 82
QY 62 LKKNPKLTLAIGGWNFGTQKFTDMVATANNROTFFVNSAIRFLRKYSFGDGLDWEYP 121
DB 83 LKKNPKLTLAIGGWNFGTQKFTDMVATANNROTFFVNSAIRFLRKYSFGDGLDWEYP 142
QY 122 GSQGSFPAVDKERFTTLVQDLANAFQEAQTSKGERLLLSAAVPAQGTYYDAGYEVDKIAQ 181
DB 143 GSQGSFPAVDKERFTTLVQDLANAFQEAQTSKGERLLLSAAVPAQGTYYDAGYEVDKIAQ 202
QY 182 NLDVNLMAIDYDFHGSWEKVTGHSNPLYKQESGAAASLNVDAAVOQMLQKGTTPASKLI 241
DB 203 YLDYTHVMTYDLHGSWEKVTGHSNPLYKQESGAAASLNVDAAVOQMLQKGTTPASKLI 262
QY 242 GMPTTGRSFTLASSSDTRVGA PATGSGTGPPTKGGMLAYEVCWS--KGATKQRIODQ 299
DB 263 GMPTTGRSFTLASSSDTRVGA PATGSGTGPPTKGGMLAYEVCWS--KGATKQRIODQ 322
QY 300 KVPYIFRDQWVGDDVESFKTKVSYLKQKGLGGAMWALDLDLDPAGFSCNQGRIPIOT 359
DB 323 EVPYAYQGNVWVGVDNIKSFDIKAQWLKHNFKGAMVWALDLDLDTGTTCNOCKFPLIST 382
QY 360 LRQELSLPVLPSGTPELEVPKGPQSEP-----EHGPPSGQDTFCQKAD 404
DB 383 LRKAL-----GLOSASCTAPAQPIEPITAAPSGSGSGSGSGSGSGGFCFAVRAN 435
QY 405 GLYPNPRSSFYSCAAGRLFOQSCPTGLVFNNSCKCCTW 444
DB 436 GLYPVANNRNFHWCVNGVTTYQONCOAGLVFDTSCDCCNW 475

RESULT 4

```

US-10-004-219B-9
; Sequence 9, Application US/10004219B
; Publication No. US20030087414A1
; GENERAL INFORMATION:
; APPLICANT: Macrozyme
; APPLICANT: Berts, Johannes M.F.G.
; APPLICANT: Boot, Rolf G.
; TITLE OF INVENTION: A mammalian mucinase, its recombinant production, and
; TITLE OF INVENTION: its use in therapy or prophylaxis against diseases in
; TITLE OF INVENTION: which mucus is involved or infection diseases
; FILE REFERENCE: 2183-5136US
; CURRENT APPLICATION NUMBER: US/10/004,219B
; CURRENT FILING DATE: 2001-11-02
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 9
; LENGTH: 452
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1)..(452)
; OTHER INFORMATION: /note="Mouse AMCase"
US-10-004-219B-9

```

TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
FILE REFERENCE: MEI-030
CURRENT APPLICATION NUMBER: US/10/097,340
CURRENT FILING DATE: 2002-03-14
PRIOR APPLICATION NUMBER: 60/276,025
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/325,149
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 60/276,026
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/324,967
PRIOR FILING DATE: 2001/09/26
PRIOR APPLICATION NUMBER: 60/311,732
PRIOR FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: 60/325,102
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 60/323,590
PRIOR FILING DATE: 2001-09-19
NUMBER OF SEQ ID NOS: 363
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 45
LENGTH: 383
TYPE: PRT
ORGANISM: Homo sapiens
US-10-097-340-45

Query Match 43.4%; Score 1041.5; DB 9; Length 383;
Best Local Similarity 53.3%; Pred. No. 4.5e-83;
Matches 195; Conservative 66; Mismatches 96; Indels 9; Gaps 5;
QY 2 KLVCFYFNWQAQYRQGEAREPLPKDLPDLPSLCTHLIYAFAGTNNHOLSTTEWNETLYQEFNG 61
DB 23 KLVCFYISWSQYREGDSCFPDLDRLFLCTHLIYSPANISNDHIDTWENDVTLXGMLNT 82
QY 62 LKKNPKLKTLLAIGGNWFGTQKFTDMVATANNROTFFVNSAIRFLKYSDFDGLDWEYP 121
DB 83 LKKNPNLKTLLSVGGWNGFSQRFKSIASNTQSRRTFKSVPPFLRTHGDFGLDLAWLYP 142
QY 122 GSQGSFPAVDKERFTTLVODLANAFOQEAQTSKGERLLLSAAVPAQOTYVDAGYVDKIAQ 181
DB 143 GR-----DKQHFTTLIKEMKAEFKEAQ-PGKQKLLLSAALSAGKVTIDSSYDIKISQ 196
QY 182 NLDVFNLMAYDFHGSWEKVTGHNSPLYKQESGAAASLVNDAVQOVLQKGTGPASKLIL 241
DB 197 HLDIFSILTYDFHGAWRQTTHHSPLFRGQGDASDRFSNADYAVSVYVLRGAPANKLYM 256
QY 242 GMPTVGRSFTLASSSDTRVGAPATSGTPPGFTKEGGMLAYEYVCSW-KGATKQRIQDQK 300
DB 257 GIPTFGRSFTLA-SSETGVGAPISGPGIPGRFTKEAGTLAYEYICDPLRGATVHRTLGQQ 315
QY 301 VPYIFRDNQWGFDDVESFPTKVSYLKQKLGAMVWALDLDLDPFQSGFCQDLRPLTNA 375
DB 360 LRQELS 365
DB 376 IKDALA 381

RESULT 7
US-09-459-749D-17
Sequence 17, Application US/09459749D
Patent No. US20020136716A1
GENERAL INFORMATION:
APPLICANT: Millis, Albert J. T.
TITLE OF INVENTION: Compositions and Methods For Altering Cell Migration
FILE REFERENCE: 0794.016A
CURRENT APPLICATION NUMBER: US/09/459,749D
CURRENT FILING DATE: 1999-12-10
PRIOR APPLICATION NUMBER: 60/111,856
PRIOR FILING DATE: 1998-12-11
NUMBER OF SEQ ID NOS: 17

SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 17
LENGTH: 383
TYPE: PRT
ORGANISM: Sus scrofa
US-09-459-749D-17
Query Match 42.6%; Score 1021.5; DB 10; Length 383;
Best Local Similarity 52.2%; Pred. No. 2.6e-81;
Matches 191; Conservative 67; Mismatches 99; Indels 9; Gaps 5;
QY 2 KLVCFYFNWQAQYRQGEAREPLPKDLPDLPSLCTHLIYAFAGTNNHOLSTTEWNETLYQEFNG 61
DB 23 KLVCFYISWSQYREGDSCFPDLDRLFLCTHLIYSPANISNDHIDTWENDVTLXGMLNT 82
QY 62 LKKNPKLKTLLAIGGNWFGTQKFTDMVATANNROTFFVNSAIRFLKYSDFDGLDWEYP 121
DB 83 LKKNPNLKTLLSVGGWNGFSQRFKSIASNTQSRRTFKSVPPFLRTHGDFGLDLAWLYP 142
QY 122 GSQGSFPAVDKERFTTLVODLANAFOQEAQTSKGERLLLSAAVPAQOTYVDAGYVDKIAQ 181
DB 143 GR-----DKRHFTTLIKEMKAEFKEAQ-LPQTERLLLSGAVSAGKVAIDRGYDIAQISQ 196
QY 182 NLDVFNLMAYDFHGSWEKVTGHNSPLYKQESGAAASLVNDAVQOVLQKGTGPASKLIL 241
DB 197 HLDIFSILTYDFHGAWRQTTHHSPLFRGQGDASDRFSNADYAVSVYVLRGAPANKLYM 256
QY 242 GMPTVGRSFTLASSSDTRVGAPATSGTPPGFTKEGGMLAYEYVCSW-KGATKQRIQDQK 300
DB 257 GIPTFGRSFTLASSK-TDVGAPASGPGIPGRFTKEGILAYEYICDPLQCATVRRPLGQQ 315
QY 301 VPYIFRDNQWGFDDVESFPTKVSYLKQKLGAMVWALDLDLDPFQSGFCQDLRPLTSA 375
DB 360 LRQELS 365
DB 376 IKDOLA 381
RESULT 8
US-10-218-743-41
Sequence 41, Application US/10218743
Publication No. US20030096779A1
GENERAL INFORMATION:
APPLICANT: McCall, Catherine A.
APPLICANT: Hunter, Shirley Wu
TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
TITLE OF INVENTION: AND USES THEREOF
FILE REFERENCE: AL-2-C3
CURRENT APPLICATION NUMBER: US/10/218,743
CURRENT FILING DATE: 2002-08-13
PRIOR APPLICATION NUMBER: US/09/292,225
PRIOR FILING DATE: 1999-04-15
PRIOR APPLICATION NUMBER: 60/098,909
PRIOR FILING DATE: 1998-09-02
PRIOR APPLICATION NUMBER: 60/085,295
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/098,565
PRIOR FILING DATE: 1998-04-17
PRIOR APPLICATION NUMBER: 09/062,013
NUMBER OF SEQ ID NOS: 49
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 41
LENGTH: 490
TYPE: PRT
ORGANISM: Dermatophagoides farinae
US-10-218-743-41
Query Match 31.8%; Score 762; DB 9; Length 490;
Best Local Similarity 33.3%; Pred. No. 2.2e-58;

Db 93 -EKGVERFNNLRKNPELTTMISLGGWYEGSEKYSMDAANFYRQQFTQSVDLFLQYK 151
Qy 111 FDGLDLDWEYPSQ-GSPAVDKERFTLLVQDLANAFQBEAQTSGRERLLLSAAVPAQTY 169
Db 152 FDGLDLDWEYPSQRLGNPKIDKQNYLALVRELKDAEPHG-----YLLTAASFGDK 204
Qy 170 VDAGYEVKIAQNLDVFNLMAYDFHGSWEKVTGHNSPLYKROESGAAS-LNVDAAVQ 228
Db 205 IDRAVDIKELNKLFDMMNMTYDYGWENFYGHNAFLYKRPDETDLHTYFNVNTHY 264
Qy 229 WLOKTPASKILGMPTYGRSFTTLASSSDTRVGAPATGSGTPGPTKEGGMAYVEVCSW 288
Db 265 YLNNGATDKLVGMVFPFYGRAMSIEDRSKLGDPKAGMSPGFTSGEGVLSYIELCOL 324
Qy 289 KGATKQRIQ-DQ--KVPIFRDNQWGFDDVESFKTKVSYLKQKGLGAMVWALDDFA 345
Db 325 FQKEWHIQYDEYNAPYGYNDKIWGVYDDLASISCKLAFLKELGVSVMVWLENDDPK 384
Qy 346 G-----FSCNQG-----RYPLIOTLROELSPLYPST 373
Db 385 GHCGPKNLLNKVHNMINGDEKNSPECILGSPSTTTPTTTPTTTPTTPTTPTTPTT 444
Qy 374 PELEVP-KPGQSEPEHGPSGQDT 397
Db 445 PSPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 469

RESULT 13

US-10-218-743-18
; Sequence 18, Application US/10218743
; Publication No. US20030096779A1
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
; AND USES THEREOF
; FILE REFERENCE: AL-2-C3
; CURRENT APPLICATION NUMBER: US/10/218,743
; PRIOR FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: US/09/292,225
; PRIOR FILING DATE: 1999-04-15
; PRIOR APPLICATION NUMBER: 60/098,909
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/085,295
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/098,565
; PRIOR FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: 09/062,013
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 18
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Dermatophagoides farinae
US-10-218-743-18

Query Match 30.8%; Score 738; DB 9; Length 555;
Best Local Similarity 34.2%; Pred. No. 3.4e-56;
Matches 152; Conservative 85; Mismatches 150; Indels 58; Gaps 11;
Qy 2 KLVCYFTNWAQYRQGEARFLPKDLPSCLTLLIYAFAGMTNHLST-----TEW 50
Db 34 RIVCYVGTWSVTHKDDP-YTIEDIDPFKCTHLMYGFADKDEYKTIQVDFPDYQDDNHSN 92
Qy 51 NDELYQEFNGLKMNPKLKTLLAIGGNFQTOKFTDMVATANNRQTFVNSAIRFLKYS 110
Db 93 -EKGVERFNNLRKNPELTTMISLGGWYEGSEKYSMDAANFYRQQFTQSVDLFLQYK 151
Qy 111 FDGLDLDWEYPSQ-GSPAVDKERFTLLVQDLANAFQBEAQTSGRERLLLSAAVPAQTY 169
Db 152 FDGLDLDWEYPSQRLGNPKIDKQNYLALVRELKDAEPHG-----YLLTAASFGDK 204

Qy 170 VDAGYEVKIAQNLDVFNLMAYDFHGSWEKVTGHNSPLYKROESGAAS-LNVDAAVQ 228
Db 205 IDRAVDIKELNKLFDMMNMTYDYGWENFYGHNAFLYKRPDETDLHTYFNVNTHY 264
Qy 229 WLOKTPASKILGMPTYGRSFTTLASSSDTRVGAPATGSGTPGPTKEGGMAYVEVCSW 288
Db 265 YLNNGATDKLVGMVFPFYGRAMSIEDRSKLGDPKAGMSPGFTSGEGVLSYIELCOL 324
Qy 289 KGATKQRIQ-DQ--KVPIFRDNQWGFDDVESFKTKVSYLKQKGLGAMVWALDDFA 345
Db 325 FQKEWHIQYDEYNAPYGYNDKIWGVYDDLASISCKLAFLKELGVSVMVWLENDDPK 384
Qy 346 G-----FSCNQG-----RYPLIOTLROELSPLYPST 373
Db 385 GHCGPKNLLNKVHNMINGDEKNSPECILGSPSTTTPTTTPTTPTTPTTPTTPTT 444
Qy 374 PELEVP-KPGQSEPEHGPSGQDT 397
Db 445 PSPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 469

RESULT 14

US-09-923-844B-2
; Sequence 2, Application US/09923844B
; Patent No. US20020166143A1
; GENERAL INFORMATION:
; APPLICANT: Pioneer Hi-Bred International, Inc.
; APPLICANT: Bao, Zhongmeng
; APPLICANT: Lu, Guohua
; TITLE OF INVENTION: Sclerotinia-inducible Genes and
; Promoters and Their Uses
; FILE REFERENCE: 35718/234631
; CURRENT APPLICATION NUMBER: US/09/923,844B
; CURRENT FILING DATE: 2001-08-07
; PRIOR APPLICATION NUMBER: US 60/224,603
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 371
; TYPE: PRT
; ORGANISM: Helianthus annuus
US-09-923-844B-2

Query Match 20.3%; Score 487; DB 9; Length 371;
Best Local Similarity 36.1%; Pred. No. 2e-34;
Matches 126; Conservative 44; Mismatches 135; Indels 44; Gaps 15;
Qy 6 YFTNWAQYRQGEARFLPKDLPSCLTLLIYAFAGMTN-----HQLSTTWNDETLYQ 57
Db 30 YWPSWAQ-----DFLPSNIQTAYFTHVYAFSLPNNVTFQFVHRTTASALN----- 77
Qy 58 EFN-GLKQNNPKLKTLLAIGGNFQTOKFTDMVATANNRQTFVNSAIRFLKYSFDGLD 115
Db 78 SFNTALHGKNPPVKTLFSGISGAGVKQLFSLASSPGSRAAFIRSTIQVARNYFEDGAD 137
Qy 116 LDWEYPSGSGSPAVDKERFTTLVQDLANAFQBEAQTSGRERLLLSAAVPAQTYVDAG-- 173
Db 138 LDWEYPETQ-----TDMNFGULLDEWRVAVNNEATSTGKPRLLLSAATRHBEVRDNGVA 193
Qy 174 -YEVDKIAQNLDVFNLMAYDFHGSW-EKVGTGHNSPLYKROESGAAS-LNVDAAVQ 231
Db 194 KYPVASINKNLGDNAMCYDHGTPPTPATGAPALYNPN-----GSLSTNSGLQSMIS 247
Qy 232 KGTASKILGMPTYGRSFTTLASSSDTRVGAPATGSGTPGPTKEGGMAYVEVCSW 291
Db 248 AGIQOKLVGMVPLYGWTWKLKNPSVINGIGAPAAIG-PG---NEGAML-YSEVQOQFNAQ 302
Qy 292 TKQRI--ODQKVPYI-FRDNQWGFDDVESFKTKVSYLKQKGLGAMVW 337
Db 303 NNARVYDTQTVSYYSYSTGTTWIGYDDVNSVQRKQYAKSLNIGGYFFW 351

RESULT 15
US-09-748-033-3
; Sequence 3, Application US/09748033
; Patent No. US20020069431A1
; GENERAL INFORMATION:
; APPLICANT: Broadway, Roxanne M.
; TITLE OF INVENTION: EFFECT OF ENDOCHITINASE AND CHITOSIDASE AND THEIR
; FILE REFERENCE: 19603/3091
; CURRENT APPLICATION NUMBER: US/09/748,033
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/172,003
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 3
; LENGTH: 376
; TYPE: PRT
; ORGANISM: Streptomyces albidoflavus
US-09-748-033-3

Query Match : 14.0%; Score 335.5; DB 10; Length 376;
Best Local Similarity 29.0%; Pred. No. 4e-21;
Matches 116; Conservative 43; Mismatches 134; Indels 107; Gaps 19;

Qy	6	YFTNWAQY-RQGEARFLPKDLSLCTHLIYAFAGTNHQLSTTE-----	49
Db	14	YFTEWGVGNVHVKNLVTSGSAEKITHINYSFGVGGKCTIGDSFAAYDKAYTAESV	73
Qy	50	-----WNDETLYQEFNGLKKNM---PKLKLIALIGWNFGTQKFTDMVATANNRQTFVN	100
Db	74	DGVADTW-DQPLRGNFNQLEKAKYPHIKVLWSFGWTW---SGFTDAV---KNPAFAK	128
Qy	101	SAIRFLR--KYS--FDGLDLWYPSQG-----SPVDKERFTTLVQ-----DLA	142
Db	129	SCHDLVEDPRWADVDFDGLDLWYFNACGLSCDSGPAALK-----NMVQAMRAQFGTDLV	184
Qy	143	N-AFQOEAOQTSGRERLLLSAAVPAGQTYVDAGYEVDKIAQNLDVNLMAVDHGSWEKVT	201
Db	185	TAAITADASSGGK-----LDAAAYAG-----AAQYFDWYNVMTYDFGAWDK-T	227
Qy	202	GHNSPLYKROESG-AAASLNVDAAVQWMLQKTPASKLILGMPTYGRSFTLASSSDTRV	260
Db	228	GPTAPHSAIISYSGIPKADPHSAATAKAKAGVPASKLLIGIGFYGRW-----TGV	280
Qy	261	GAPATSGTGPPTKEGMLAYYEVCWSKGATKQRTODOKV-----PYIFR	306
Db	281	TODAPGCTATGPAT-----GTAEAGIEDYKVLKNTCPATGTVGGTAYAKC	325
Qy	307	DNQWVGFDVDESFKTVSYLKKGLGAMWALDLDLDDFAG	346
Db	326	GSNWSYDTPATIKTKMTWAKOGLGAFWFESGDTAGG	365

Search completed: June 29, 2003, 21:13:35
Job time : 19.1154 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: June 29, 2003, 21:03:33 ; Search time 2892.5 Seconds
(without alignments)
4477.355 Million cell updates/sec

Title: US-10-004-219b-10
Perfect score: 2398
Sequence: 1 AKLCVFTNWAQYRGARF.....QSCPTGLVFSNCKRCTWN 445

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DEV=xl

-MODEL=frame_p2n.model -DEV=xl
-Q/cgn2_1/USPTO.spool/US10004219/runat_24062003_160228_8159/app.query.fasta_1.1877
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human4 -cdi -LINT=45
-DOCCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -SCORE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10004219 @CGN 1.1 6828 @runat_24062003_160228_8159 -NCPU=6 -ICPU=3
-NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOC=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl :
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.ev.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2398	100.0	1633	9 HSU29615	U29615 Human chito
2	2398	100.0	1636	6 AR206041	AR206041 Sequence
3	2398	100.0	1636	6 AR212138	AR212138 Sequence
4	2398	100.0	1636	6 AX108750	AX108750 Sequence
5	2394	99.8	1643	6 AR172107	AR172107 Sequence
6	2392	99.7	1656	6 AR206042	AR206042 Sequence
7	2392	99.7	1656	6 AR212139	AR212139 Sequence
8	2392	99.7	1656	6 AX108752	AX108752 Sequence
9	2352.5	98.1	1710	9 HSU62662	U62662 Homo sapien
10	2348.5	97.9	1713	6 AR172108	AR172108 Sequence
11	2270.5	94.7	1599	9 AK055165	AK055165 Homo sapi
12	1323	55.2	1625	9 AF290004	AF290004 Homo sapi
13	1309.5	54.6	1526	10 AF154571	AF154571 Mus muscu
14	1309.5	54.6	1538	10 BC011134	BC011134 Mus muscu
15	1309.5	54.6	1557	10 BC034548	BC034548 Mus muscu
16	1304	54.4	1529	4 AB051629	AB051629 Bos tauru
17	1302.5	54.3	1562	5 BJA345054	AJ345054 Bufo japo
18	1301.5	54.3	1530	10 AF290003	AF290003 Mus muscu
19	1072.5	44.7	1369	6 AX405989	AX405989 Sequence
20	1072	44.7	1354	9 AB025008	AB025008 Homo sapi
21	1056	44.0	1801	9 BC008568	BC008568 Homo sapi
22	1053.5	43.9	1056	10 AF062038	AF062038 Rattus no
23	1051	43.8	1741	9 HUMA3G	M80927 Human glyco
24	1046.5	43.6	1433	6 AR042834	AR042834 Sequence
25	1046.5	43.6	1500	9 HSU58515	U58515 Human chiti
26	1046.5	43.6	1526	6 AR042835	AR042835 Sequence
27	1045.5	43.6	1418	9 HSU49835	U49835 Human YKL-3
28	1045.5	43.6	1434	9 HSU58514	U58514 Human chiti
29	1043.5	43.5	1149	6 E01500	E01500 cDNA encodi
30	1043.5	43.5	1149	6 E01501	E01501 cDNA encodi
31	1043	43.5	1449	9 BC011460	BC011460 Homo sapi
32	1030.5	43.0	1733	4 SSU19900	U19900 Sus scrofa
33	1030.5	43.0	1733	4 SSGP38KD	247803 S.scrofa 38
34	1029.5	42.9	1281	4 AY081150	AY081150 Capra hir
35	1015.5	42.3	3946	3 AY051988	AY051988 Drosophill
36	1007	42.0	1641	10 BC004734	BC004734 Mus muscu
37	1007	42.0	1651	10 BC003780	BC003780 Mus muscu
38	1007	42.0	1664	10 BC005611	BC005611 Mus muscu
39	1000	41.7	1616	10 MMBP39	X93035 M.musculus
40	986.5	41.1	6439	3 AB074977	AB074977 Haemaphys
41	979	40.8	997	4 AF011373	AF011373 Bos tauru
C 42	975.5	40.7	173298	2 AC006489	AC006489 Drosophill
C 43	975.5	40.7	176521	3 AC097725	AC097725 Drosophill
C 44	975.5	40.7	304633	3 AE003477	AE003477 Drosophill
45	970	40.5	1209	10 AY049765	AY049765 Mus muscu

ALIGNMENTS

RESULT 1

HSU29615
 LOCUS 1633 bp mRNA linear PRI 03-NOV-1995
 DEFINITION Human chitotriosidase precursor mRNA, complete cds.
 ACCESSION U29615
 VERSION U29615.1 GI:1050957
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1633)
 Boot.R.G., Renkema,G.H., Strijland,A., van Zonneveld,A.J. and
 Aerts,J.M.
 Cloning of a cDNA encoding chitotriosidase, a human chitinase
 produced by macrophages
 J. Biol. Chem. 270 (44), 26252-26256 (1995)
 MEDLINE 96064695
 PUBMED 7592832
 REFERENCE 2 (bases 1 to 1633)
 AUTHORS Boot,R.G.
 TITLE Direct Submission
 JOURNAL Submitted (21-JUN-1995) Rolf G. Boot, Department of Biochemistry,
 AMC, E.C. Slater, Institute, University of Amsterdam, Meibergdreef
 15, 1105 AZ, Amsterdam, The Netherlands
 FEATURES
 source
 1..1633
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="Chl.1"
 /cell_type="macrophage"
 /clone_lib="macrophage-1 cDNA library of Rolf G. Boot"
 13..1413
 /codon_start=1
 /product="chitotriosidase precursor"
 /protein_id="AAC50246.1"
 /db_xref="GI:1050958"
 /translation="MVRSAAGFMVLLIPMGSAALVCYFTNWAQYRQGEARFLPK
 DLDPSLCTLLIYAFAGMTNHLSTEWNETLYQEFNGLKWNPKLTKLLIAGWNFG
 TOKFTDMVATANNQRFVNSAIRFLRYKSFGLDLDEYVPGSGSPAVDKERFTLLVQ
 DLANAFOQAQTSQSKELLISRAVPAGQYVDAGYEVDKIAONLDFVNLMAVDFHGSW
 EKVTHNSFLYKQESGAASLVNDAVQOQLKGTSPASKILILGHPYTGRTFLASS
 SDTRVGAATQSPFPFTKEGGLAYIEVCSWKGATKQRIODQKVPYIFRDNQWGF
 DVSEFKTVSLYKQGLGAMVWALDDFAFGSCNQGRIYPLIQLRQELSLPYLPS
 GTPELEVPKPGPSPEPEHGPSQDQTFQCKADGLYPNPRERSFSYSCAAGRLFQGS
 PTGLVPSNCKCCTWN"
 13..75
 sig_peptide
 mat_peptide
 76..1410
 /product="chitotriosidase"
 /notes="human chitinase"
 polyA_signal
 polyA_site 1605..1610
 1633
 /note="10 A nucleotides"
 BASE COUNT 354 a 489 c 443 g 347 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 9,34e-176 Length: 1633
 Score: 2398.00 Matches: 445
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0
 US-10-004-219B-10 (1-445) x HSU29615 (1-1633)
 QY 1 AlaLysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPhe 20
 Db 76 GCAAAACTGCTGCTACTTCCAACTGGGCCAGTACAGACAGGGGAGGCTCGCTTC 135
 QY 21 LeuProLysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet 40
 Db 136 CTGCCCAAGAGACTTGGACCCACCCCTTTGACCCACCTATCTACGCCCTTCGCTGCATG 195

QY 41 ThrAsnHisGlnLeuSerThrThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsn 60
 Db 196 ACCAACCCACCCAGCTGAGCACCACTGAGTGAATGACGAGACTCTCTACGAGGTTCAAT 255
 QY 61 GlyLeuLysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyClyTrpAsnPhe 80
 Db 256 GGCCTGAAGAAGATGAATCCCAAGCTGAAGCCCTTTAGCCATCGGAGCTGGAAATTC 315
 QY 81 GlyThrGlnLysPheThrAspMetValAlaThrAlaAsnAsuArgGlnThrPheValAsn 100
 Db 316 GGCATCTCAGAGTTTACAGATATGTAGCCACCGCCACCAACCGTCAGACCTTTGTCAAC 375
 QY 101 SerAlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyr 120
 Db 376 TCGGCCATCAGGTTTCTCGCAATACAGCTTTGACGGCTTGACCTTGAGTGGAGTAC 435
 QY 121 ProGlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAsp 140
 Db 436 CCAGGAAGCCAGGGAGCCCTCGCTAGCAAGAGCGCTTCACAAACCTCGGTACAGGAC 495
 QY 141 LeuAlaAsnAlaPheGlnGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuSer 160
 Db 496 TTGGCCATGCTTTCAGCAGGAGCCACAGACTCAGGGAAGAACGCTTCTTCTGAGT 555
 QY 161 AlaAlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAla 180
 Db 556 GCAGCGGTTCCAGCTGGGCGAGACCTATGTGGATGCTGGATACGAGTGACAAATCGCC 615
 QY 181 GlnAsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysVal 200
 Db 616 CAGAACTCTGGATTTTGTCAACCTTATGGCTTACGACTTCCATGGCTCTTTGGGAGAGGTC 675
 QY 201 ThrGlyHisAsnSerProLeuTyrLysArgGlnGluSerGlyValAlaAlaSerLeu 220
 Db 676 ACGGACATACAGCCCTCTTACAGAGGCAAGAGAGTGTGTGCAGCAGCCAGCCTC 735
 QY 221 AsnValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuIle 240
 Db 736 AACGTGGATGCTGCTGTGCAACAGTGGCTGCAGAGGGGACCCCTGCCAGCAAGCTGATC 795
 QY 241 LeuGlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgVal 260
 Db 796 CTTGGCATGCTTACCTACGAGCGCTCTTACACCTGGCTCTCTCATCAGACACCAAGAGTG 855
 QY 261 GlyValProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyGlyMetLeu 280
 Db 856 GGGGCCCCAGCCACAGGCTCTGGCACTCCAGGCCCTTCCACAGGAGAGGAGGATGCTG 915
 QY 281 AlaTyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLys 300
 Db 916 GCCTACTATGAAGTCTGCTCTCTGGAAGGGGGCCACCAACACAGAAATCCAGGATCAGAAG 975
 QY 301 ValProTyrIlePheArgAspAsnGlnTrpValGlyPheAspAspValGluSerPheLys 320
 Db 976 GTGCCCTATCATCTCCGGGACCAACAGTGGTGGGCTTTGATGATGTGGAGAGCTTCAAA 1035
 QY 321 ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyGlyValAlaMetValTrpAlaLeuAsp 340
 Db 1036 ACCAAGGTGAGTATCTGAGCAGAGAGGAGCTGGGGGGGCCCATGTCTGGGCATCGAC 1095
 QY 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeu 360
 Db 1096 TTAGATGACTTTGGCGGCTTCTCTGCAACAGGCGCGATACCCCTCATCCAGACGCTA 1155
 QY 361 ArgGlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLys 380
 Db 1156 CGGCAAGAACTGAGTCTTCCATCTTGGCTTTCAGGACCCCGAGAGCTTGAAGTTCACAAA 1215
 QY 381 ProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGln 400
 Db 1216 CCAGGTGAGCCCTCTGAACTGAGCATGGCCCGCCCTGGCAAGACACGCTTCTGCCAG 1275
 QY 401 GlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyrSerCysAla 420

Db 1276 GCGAAAGCTGATGGGCTCTATCCCAATCTCGGAAACGGTCCAGCTTCTACAGCTGTGCA 1335
Qy 421 AlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerHsnSerCysLys 440
Db 1336 GCGGGGCGGCTGTTCAGCAAAAGCTGCCCGACAGGCGCTGGTGTTCAGCAACCTCTGCAAA 1395
Qy 441 CysCysThrTirAsn 445
Db 1396 TGCTGCACCTGGAAAT 1410
RESULT 2
AR206041
LOCUS AR206041 1636 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 1 from patent US 6372212.
ACCESSION AR206041
VERSION AR206041.1 GI:21504525
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1636)
AUTHORS Gray,P.W.
TITLE Chitinase materials and methods
JOURNAL Patent: US 6372212-A 1 16-APR-2002;
FEATURES
Location/Qualifiers
1..1636
/organism="unknown"
BASE COUNT 361 a 491 c 440 g 344 t
ORIGIN
Alignment Scores:
Pred. No.: 9.36e-176 Length: 1636
Score: 2298.00 Matches: 445
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-004-219b-10 (1-445) x AR206041 (1-1636)
Qy 1 AlaLysLeuValCysTyrPheThrAsnTirAlaGlnTyrArgGlnGlyGluAlaArgPhe 20
Db 65 GCAAAACTGGTCTGTCTACTTACCAACTGGGCCAGTACAGACAGGGGGAGGCTCGCTTC 124
Qy 21 LeuProLysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet 40
Db 125 CTGCCCAAGGACTTGGACCCCGCCAGCTTTGCACCCACCTCATCTACCGCTTCGCTGGCATG 184
Qy 41 ThrAsnHisGlnLeuSerThrThrGluTirPAsnAspGluThrLeuTyrGlnGluPheAsn 60
Db 185 ACCAACCCAGCTGAGCACCCTAGTGGATGACGAGACTCTCTACCGAGGATTCAT 244
Qy 61 GlyLeuLysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTirAsnPhe 80
Db 245 GGCTCTGAAGAAGATGAATCCCAAGCTGAAGACCCCTGTAGCCATCGAGGCTGGAATTC 304
Qy 81 GlyThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsn 100
Db 305 GGCACCTCAGAAAGTTTACAGATATGGTAGCCACCGCCCAACACCGTTCAGACCTTTGTCAAC 364
Qy 101 SerAlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTirGluTyr 120
Db 365 TCGGCCATCATGGTTTCTGCGCAATATACAGCTTTGACGGCTTGACCTTGACTGGAGTAC 424
Qy 121 ProGlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAsp 140
Db 425 CCAGGAAGCCAGGGGAGCCCTGCGGTAGACAGAGCGCTTCAACACCCCTGGTACAGGAC 484
Qy 141 LeuAlaAsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuSer 160
Db 485 TTGGCCAATCCCTTCCAGCAGGAAGCCAGACCTCAGGGAAGGAGCGCTTCTTCTGAGT 544

Qy 161 AlaAlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAla 180
Db 545 GCAGCGGTTCCAGCTGGCGACAGACCTATGTGATGTGATGACGAGGTGACAAAAATCGCC 604
Qy 181 GlnAsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTirGluLysVal 200
Db 605 CAGAAACCTGGATTTTGTCAACCTTATGGCTTACGACTTCCATGGCTCTTGGGAGAGGTC 664
Qy 201 ThrGlyHisAsnSerProLeuTyrLysArgGlnGluSerGlyValAlaAlaSerLeu 220
Db 665 ACGGACATACAGCCCTCTACAGAGGACAGAGAGAGTGGTGCAGCAGCCAGCCTC 724
Qy 221 AsnValAspAlaAlaValGlnGlnTirPLeuGlnLysGlyThrProAlaSerLysLeuIle 240
Db 725 AACGTGGATGTCTGTGCAACAGTGGCTGCAGAGGGGACCCCTGCCAGCAAGCTGATC 784
Qy 241 LeuGlyMetProThrTyrGlyArgSerPheThrIleuAlaSerSerSerAspThrArgVal 260
Db 785 CTTGGCATGCCCTTACCTACGGACGCTCTTCACACTGGGCTCTCATCAGACACCCAGAGTG 844
Qy 261 GlyAlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyGlyMetLeu 280
Db 845 GGGGCCCCAGCCACAGGGTCTGGCACTCCAGGCCCTTCCACCAAGGAGGAGGATGCTG 904
Qy 281 AlaTyrTyrGluValCysSerTirPlysGlyAlaThrLysGlnArgIleGlnAspGlnLys 300
Db 905 GCCTACTATGAAGTCTGCTCTCTGGAAGGGGGCCACCAACAGAGAATCCAGGATCAGAAG 964
Qy 301 ValProTyrIlePheArgAspAsnGlnTirPValGlyPheAspAspValGluSerPheLys 320
Db 965 GTGCCCTACATCTCCGGGACCAACAGTGGGTGGGCTTTGATGATGTGGAGAGCTTCAAA 1024
Qy 321 ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyGlyAlaMetValTirPAlaLeuAsp 340
Db 1025 ACCAAGCTCAGCTATCTGACGACGAGGAGCTGGGGGGCCCATGGTCTGGGCACTGGAC 1084
Qy 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeu 360
Db 1085 TTAGATGACTTTTGGCGGCTTCTCTCTCAACAGGGGCCGATACCCCTCATCCAGACGCTA 1144
Qy 361 ArgGlnGlnLeuSerLeuProTyrLeuProSerGlyThrProGluLeuValProLys 380
Db 1145 CGGCAAGGAACTCAGTCTTCCATACCTTGGCTTCCAGCACCCCGAGCTTGAAGTTCAAAA 1204
Qy 381 ProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGln 400
Db 1205 CCAGGTGACCCCTCTGACCTGAGCATGGCCCGCCCGCCCTGGACACACGTTCTGCGCAG 1264
Qy 401 GlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyrSerCysAla 420
Db 1265 GGCAAAGCTGATGGGCTCTATCCCAATCTCGGGAACGGTCCAGCTTCTACAGCTGTGCA 1324
Qy 421 AlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerHsnSerCysLys 440
Db 1325 GCGGGGCGGCTTTCAGCAAAAGCTGCCCGACAGGCGCTGGTGTTCAGCAACCTCTGCAAA 1384
Qy 441 CysCysThrTirAsn 445
Db 1385 TGCTGCACCTGGAAAT 1399
RESULT 3
AR212138
LOCUS AR212138 1636 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 1 from patent US 6399571.
ACCESSION AR212138
VERSION AR212138.1 GI:21515643
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1636)
AUTHORS Gray,P.W. and Tjoelker,L.W.
TITLE Chitinase chitin-binding fragments

JOURNAL Patent: US 639571-A 1 04-JUN-2002;

FEATURES

Location/Qualifiers
1. 1636
/organism="unknown"

BASE COUNT 361 a 491 c 440 g 344 t

ORIGIN

Alignment Scores:

Pred. No.:	9.36e-176	Length:	1636
Score:	2398.00	Matches:	445
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-10-004-219B-10 (1-445) x AR212138 (1-1636)

QY 1 AlaLysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyAlaAaRgPhe 20
Db 65 GCAAACTGGTCTGCTACTTACCAACTGGGCCAGGTACAGACAGGGGAGGCTCGCTTC 124
QY 21 LeuProLysAspLeuAspProSerLeuCysThrHisLeuileTyrAlaPheAlaGlyMet 40
Db 125 CTGCCCAAGGACTTGGACCCAGCCCTTTGCCACCCACCTCATCTACGCCCTTCGCTGGCATG 184
QY 41 ThrAsnHisGlnLeuSerThrThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsn 60
Db 185 ACCAACCCAGCTGAGCACCACTGAGTGGAAATGACGAGACTCTTACCAAGGATTCAT 244
QY 61 GlyLeuLysLysMetAsnProLysLeuLysThrLeuLeuAlaileGlyGlyTyrAsnPhe 80
Db 245 GGCCTGAAGAAGATGAATCCAAAGCTGAAGACCCCTGTAGCCATCGGAGGCTGGAATTC 304
QY 81 GlyThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsn 100
Db 305 GGCACCTCAGAAGTTACAGATATGGTAGCCACGCCCAACACCGTCAGACCTTTGTCAAC 364
QY 101 SerAlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuTyrGluTyr 120
Db 365 TCGGCATCAGGTTCTTGCGCAATACAGCTTTGACGGCTTGACCTTGACTGGGAGTAC 424
QY 121 ProGlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAsp 140
Db 425 CCAGGAAGCCAGGGAGCCCTGCGGTAGACAAGAGCGCTTCAACACCCCTGGTACAGGAC 484
QY 141 LeuAlaAsnAlaPheGlnGlnGlnAlaGlnThrSerGlyLysGluArgLeuLeuSer 160
Db 485 TTGGCCAAATGCCCTTCAGCAGGAGGCCAGACCTCAGGGAGGAAGACGCTTCTTCTGAGT 544
QY 161 AlaAlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAla 180
Db 545 GCAGCGGTTCCAGCTGGCAGACCTATGTGGATGCTGGATACAGGTGGACAAATCGCC 604
QY 181 GlnAsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysVal 200
Db 605 CAGAACCTGGATTTGTCAACCTTATGGCTACGACTTCCATGGCTCTTGGGAGAGGTC 664
QY 201 ThrGlyHisAsnSerProLeuTyrLysArgGlnGluSerGlyAlaAlaSerLeu 220
Db 665 ACCGGACATTAACGCCCTCTCAAGAGGCAAGAGAGAGTGGTGGCAGCGCCAGCTTC 724
QY 221 AsnValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuIle 240
Db 725 AACGTGGATGCTGCTGTGCAACAGTGGCTGCAGAGGGGACCCCTCCAGCAGCTGATC 784
QY 241 LeuGlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgVal 260
Db 785 CTTGGCATGCTACTACACGACGACCTCTTCAACAGTGGCTCTCATCAGACACAGAGTG 844
QY 261 GlyAlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyMetLeu 280
Db 845 GGGGCCCAAGCCACAGGCTTGGCACTCCAGGCGCCCTTCCCAAGGAAGGAGGATGCTG 904

QY 281 AlaTyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLys 300
Db 905 GCTACTATGAAGTCTGCTCTCTGGAAGGGGGCCACCAACACAGAGAATCCAGGATCAGAAG 964
QY 301 ValProTyrIlePheArgAspAsnGlnTrpValGlyPheAspAspValGlnSerPheLys 320
Db 965 GTGCCCTACATCTTCGGGACCAACCAAGTGGGTGGGCTTTGATGATGTGGAGAGCTTCAAA 1024
QY 321 ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyGlyAlaMetValTrpAlaLeuAsp 340
Db 1025 ACCAAGGTGAGCTATCTGAAGCAGAGGAGACTGGGGCGGCCCATGGTCTGGGCACTGGAC 1084
QY 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeu 360
Db 1085 TTAGATGACTTTGCCGGCTTCTCTGCAACACAGGGCCGATACCCCTCATCCAGAGGCTA 1144
QY 361 ArgGlnGlnLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLys 380
Db 1145 CGGCAGGAACCTGAGTCTTCATACCTTTCAGGCACCCAGAGCTTGAAGTTCCAAA 1204
QY 381 ProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGln 400
Db 1205 CCAGGTGACCCCTCTGAACCTGAGCATGGCCGCCCTGGACAAGACACGCTTCTGCCAG 1264
QY 401 GlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyrSerCysAla 420
Db 1265 GGCMAAGCTGATGGGCTCTATCCCAATCCTCGGAAACGGTCCAGCTTCTACAGCTGTGA 1324
QY 421 AlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerAsnSerCysLys 440
Db 1325 GCGGGCGGCTTCTTCAGCAAAAGCTGCCCGACAGCGCTGGTTCAGCAACCTCTGCAAA 1384
QY 441 CysCysThrTrpAsn 445
Db 1385 TGCTGCACCTGGAAT 1399

RESULT 4

LOCUS AX108750 1636 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 1 from Patent WO0123430.
ACCESSION AX108750
VERSION AX108750.1 GI:13923942
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1636)
AUTHORS Allison,D.S., Dietsch,G.N., Gray,P.W., Shaw,K.D. and Steiner,B.H.
TITLE Human chitinase immunoglobulin fusion proteins
JOURNAL Patent: WO 0123430-A 1 05-APR-2001;
ICOS CORPORATION (US)

FEATURES

source

1. 1636
/organism="Homo sapiens"
/db_xref="taxon:9606"

CDS

2. 1402
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC37767.1"
/db_xref="GI:13923943"

/translation="MVRSVAVAGFWLMLIPWGSAAKLVCYFTNWAQYRQGEARLPK
DLPSLCITLIVAFAGMTNHOLSTTEWDETLYOFNGLKKNPKLTLLAIGWNFG
TQKTFDMVATANNRQTFVNSAIRFLKYSFDGLDLWDWEYPGSQSPAVDKERFTLVQ
ELKAFQEAQTSKERLLLSAAVPAQTVYDAGVEYDKIAQNLGFVNLMAFYDFGWS
DNKTHNSPLYKROEASLNVDAVQOQLKQTPASKLILGMFTYGRSFTLAS
SDTRVGAQPTGTPGTPKEGGLAYEVCWKGATKQRIQDQKVPYIFRDNQWVG
DDVSFKTKSVLKOKGLGGAMVALDLDDPAGFSCNOGRYPLIOTLROELSPLYS
GTPELEVPKGPQSEPHGSPGQDTFCQKADGLYLPNPRSSSFYSCAAGRLFQSC
PTGLVPSNSCKCTWN"

mat_peptide

65. 1399
/product="unnamed"
BASE COUNT 361 a 491 c 440 g 344 t

ORIGIN

Alignment Scores:
Pred. No.: 9,36e-176 Length: 1636
Score: 2398.00 Matches: 445
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-004-219B-10 (1-445) x AX108750 (1-1636)

QY 1 AlaLysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPhe 20
DB 65 GCAAACTGGTCTGCTACTTACCAACTGGGCCAGTACAGACAGGGGAGGCTCGCTTC 124
QY 21 LeuProLysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet 40
DB 125 CTGCCCAAGGACTTGGACCCCAAGCCCTTTGCACCCACCTCATCTACGCCCTTCGCTGGCATG 184
QY 41 ThrAsnHisGlnLeuSerThrThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsn 60
DB 185 ACCAACCCAGCTGAGCACCACTGAGTGGAATGACGAGACTCTTACCCAGGAGTTCAAT 244
QY 61 GlyLeuLysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPhe 80
DB 245 GGCCTGAAGAAGATGAATCCCAAGCTGAAGACCCCTGTAGCCCATCGGAGGCTGGAATTTC 304
QY 81 GlyThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsn 100
DB 305 GGCACCTCAGAAGTTCACAGATATGTGTAGCCAGCGGCCAACCAACCGTCAGACCTTTGTCAAC 364
QY 101 SerAlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyr 120
DB 365 TCGGCCATCAGGTTCTTCGGCAATACACTTTCAGCGCTTTCAGCTTTCAGCTGGGAGTAC 424
QY 121 ProGlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAsp 140
DB 425 CCAGGAAGCCAGGGAGCCCTCCGTAGACAAGGAGCGCTTCACAACCTCGGTACAGGAC 484
QY 141 LeuAlaAsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuSer 160
DB 485 TTGGCCAAATGCCCTTCAGCAGGAAGCCAGACCTCAGGGAAGGAAGCGCTTCTTCTGAGT 544
QY 161 AlaAlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAla 180
DB 545 GCAGCGGTTCCAGCTGGGAGACCTATGTGGATGTGGATAGCAGGTGACAAATCGCC 604
QY 181 GlnAsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysVal 200
DB 605 CAGAACCTGGATTTGTCAACCTTATGGCTACGACTTCATGGCTTTGGGAGAAGGTC 664
QY 201 ThrGlyHisAsnSerProLeuTyrLysArgGlnGluSerGlyAlaAlaSerLeu 220
DB 665 ACGGGACATAACAGCCCTCTACAAAGAGCAAGAGAGAGTGTGCAGCAGCAGCCTC 724
QY 221 AsnValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuIle 240
DB 725 AACGTGGATGCTGCTGTGAACAGTGGCTGCAAGAGGGACCCCTCCAGCAGCTGATC 784
QY 241 LeuGlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgVal 260
DB 785 CTTGGCATGCTACCTACGGACGCTCTTTCACACTGGCCCTTCATCATCAGACACAGAGTG 844
QY 261 GlyAlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyGlyMetLeu 280
DB 845 GGGGGCCCCAGCCAGGGCTGGCACTCCAGGCCCTTCACCAAGGAAGAGGAGTGTG 904
QY 281 AlaTyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLys 300
DB 905 GCGTACTATGAAGTCTGCTCTGGGAAGGGGGGCCACCAACAGAGAATCAGGATCAGAAG 964
QY 301 ValProTyrIlePheArgAspAsnGlnTrpValGlyPheAspValGluSerPheLys 320

DB 965 GTGCCCTACATCTTCCGGGCAACCCAGTGGTGGGCTTTGATGATGTGGAGAGCTTCAAA 1024
QY 321 ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyAlaMetValTrpAlaLeuAsp 340
DB 1025 ACCAAGGTGAGCTATCTGAAGCAAGAGGAGCTGGGGGGGGCATGGTCTGGGCACTGGAC 1084
QY 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeu 360
DB 1085 TTAGATGACTTTGCGGCTTCTCTGCAACCCAGGGCCGATACCCCTCATCCAGAGCTA 1144
QY 361 ArgGlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLys 380
DB 1145 CGCAGGAACTCAGTCTTCCATCTTGCCTTTCAGGACCCAGAGCTTGAAGTTCAAAA 1204
QY 381 ProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGln 400
DB 1205 CCAGGTACAGCCCTTGAACCTGAGCATGGCCCGCCCTGGACAAGACACGTTCTGCCAG 1264
QY 401 GlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyrSerCysAla 420
DB 1265 GCAAAAGCTGATGGCTCTATCCCAATCTCGGGAACGGTCCAGCTTCTACAGCTGTGA 1324
QY 421 AlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerAsnSerCysLys 440
DB 1325 GCGGGCGGCTGTTCCAGAAAGCTGCCCGACAGGCTGGTGTTCAGCAACTCTCTGCAAA 1384
QY 441 CysCysThrTrpAsn 445
DB 1385 TGCTGCACCTGGAAT 1399
RESULT 5
AR172107 1643 bp DNA linear PAT 17-DEC-2001
LOCUS Sequence 3 from patent US 6303118.
DEFINITION AR172107
ACCESSION AR172107
VERSION AR172107.1 GI:17911598
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 1643)
Aerts,J.Maria.Francisus.Gerardus.
AUTHORS Human chitinase, its recombinant production, its use for
TITLE decomposing chitin, its use in therapy or prophylaxis against
infection diseases
JOURNAL Patent: US 6303118-A 3 16-OCT-2001;
FEATURES Location/Qualifiers
source 1..1643
BASE COUNT 364 a 490 C 442 G 347 T
ORIGIN
Alignment Scores:
Pred. No.: 1.91e-175 Length: 1643
Score: 2394.00 Matches: 444
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.83% Indels: 0
DB: 6 Gaps: 0
US-10-004-219B-10 (1-445) x AR172107 (1-1643)
QY 2 LysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPheLeu 21
DB 79 AAACCTGGTCTGCTACTTCAACCACTGGGCCAGTACAGACAGGGGAGGCTCGCTTCCTG 138
QY 22 ProLysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMetThr 41
DB 139 CCAAGGAGCTTGGACCCCGCCCTTTCAGCCACCTCATCTACGCCCTTCGCTGGCATGAC 198
QY 42 AsnHisGlnLeuSerThrThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsnGly 61

199 AACCCACGAGTGGACACCACTGAGTGGAAATGACAGAGACTCTCTACCGAGGAGTTCAATGGC 358
62 LeuLysMetAsnProLysLeuLysThrLeuAlaTleGlyGlyTrpAsnPheGly 81
259 CTGAAGAAGATGAATCCCAAGCTGAGAGACCTTTAGCCATCGAGGCTGGAAATTCGGC 318
82 ThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsnSer 101
319 ACTCAGAAGTTACAGATATGTTAGCCACCGGCCCAACACCGTCAGACCTTTGTCAACTCG 378
102 AlaIleArgPheLeuArgLysThrSerPheAspGlyLeuAspLeuAspTrpGluTrpPro 121
379 GCCATCAGGTTTCTGGCAATACAGCTTTGACGGCTTGACCTTGAGCTGGGAGTACCCA 438
122 GlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAspLeu 141
439 GGAGCGAGGGAGCCCTGCGGTAGACAGAGCGCTTCAACACCTGTTGACAGGACTTG 498
142 AlaAsnAlaPheGlnGlnGlnAlaGlnThrSerGlyLysGluArgLeuLeuSerAla 161
499 GCCAATGCTTCCAGCAGGAGGCCAGACCTCAGGGAAGGACGCTTCTTCTGAGTGCA 558
162 AlaValProAlaGlyGlnThrThrValAspAlaGlyTrpGluValAspLysIleAlaGln 181
559 GCGGTTCCAGCTGGGAGACCTATGTGGATGCTGGATACGAGGTGGACAAATCGCCAG 618
182 AsnLeuAspPheValAsnLeuMetAlaTrpAspPheHisGlySerThrGluLysValThr 201
619 AACCTGGATTTTGTCAACCTTATGGCTTACGACTTCATGGCTTCTGGGAGAGGTCACG 678
202 GlyHisAsnSerProLeuTrpLysArgGlnGlnGluSerGlyAlaAlaAlaSerLeuAsn 221
679 GGACATAACAGCCCTCTACAGAGGCAAGAGAGAGTGGTGCAGCAGCCAGCTCAAC 738
222 ValAspAlaAlaValGlnGlnThrLeuGlnLysGlyThrProAlaSerLysLeuIleLeu 241
739 GTGGATGCTGCTGTGCAACAGTGGCTGCAGAGGGGACCCCTGCCAGCAAGCTGATCCTT 798
242 GlyMetProThrTrpGlyArgSerPheThrLeuAlaSerSerAspThrArgValGly 261
799 GGCATGCTTACCTACGAGAGCTCTCTCACTGAGCTTCTCATCAGACACAGAGTGGGG 858
262 AlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyMetLeuAla 281
859 GCCCCAGCCACAGGGTGGCACTCCAGGCCCTTCAACAAAGGAGGAGGATGCTGGCC 918
282 TyrTrpGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLysVal 301
919 TACTATGAAGTCTGCTCCTGGAAAGGGGGCCACCAACAGAGATCCAGGATCAGAAAGTG 978
302 ProTyrIlePheArgAspAsnGlnTrpValGlyPheAspAspValGluSerPheLysThr 321
979 CCTACATCTTCCGGGACCAACAGTGGTGGGCTTTGATGATGGAGAGCTTCAAAACC 1038
322 LysValSerTrpLysGlnLysGlyLeuGlyGlyAlaMetValTrpAlaLeuAspLeu 341
1039 AAGCTCAGCTATCTGAAGCAGAGGAGCTGGGGGGGCCATGCTCTGGCAGCTGGACTTA 1098
342 AspAspPheAlaGlyPheSerCysAsnGlnGlyArgTrpProIleGlnThrLeuArg 361
1099 GATGACTTTTGGGGCTTCTCTGCAACCAAGGGCCGATACCCCTCATCAGAGCTTACGG 1158
362 GlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGlnValProLysPro 381
1159 CAGGAATGATCTTCCATACTTTCCTTTCAGGACCCCGAGCTTGAAGTTCCAAACCA 1218
382 GlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGlnGly 401
1219 GGTGAGCCCTCTGAACCTGAGCATGGCCCGCCAGCCCTGGCAACAGACAGCTTCTGCCAGGC 1278
402 LysAlaAspGlyLeuTrpProAsnProArgGluArgSerSerPheTrpSerCysAlaAla 421
1279 AAAGCTGATGGGCTCTATCCCAATCTCTGGGAACGGTCCAGCTTCTACAGCTGTGCAGCG 1338

QY 422 GlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerAsnSerCysLysCys 441
Db 1339 GGGCGGCTGTTCCAGCAAGAGCTGCCGACAGGCTGGTGTTCAGCAACTCTCTGCAATGC 1398
QY 442 CysThrTrpAsn 445
Db 1399 TGCACCTGGAAT 1410
RESULT 6
AR206042 AR206042 1656 bp DNA linear PAT 20-JUN-2002
LOCUS Sequence 3 from patent US 6372212.
DEFINITION AR206042
ACCESSION AR206042
VERSION AR206042.1 GI:21504526
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1656)
AUTHORS Gray, P.W.
TITLE Chitinase materials and methods
JOURNAL Patent: US 6372212-A 3 16-APR-2002;
FEATURES Location/Qualifiers
1..1656
/organism="unknown"
BASE COUNT 365 a 497 c 447 g 347 t
ORIGIN
Alignment Scores:
Pred. No.: 2,76e-175 Length: 1656
Score: 2392.00 Matches: 444
Percent Similarity: 99.78% Conservative: 0
Best Local Similarity: 99.78% Mismatches: 1
Query Match: 99.75% Indels: 0
DB: 6 Gaps: 0
US-10-004-219B-10 (1-445) x AR206042 (1-1656)
QY 1 AlaLysLeuValCysThrPheThrAsnTrpAlaGlnTrpArgGlnGlyGluAlaArgPhe 20
Db 90 GCAAACTGGTCTGCTACTTACCACCTGGGCCAGTACAGACAGGGGAGCTCGCTTC 149
QY 21 LeuProLysAspLeuAspProSerLeuCysThrHisLeuIleTrpAlaPheAlaGlyMet 40
Db 150 CTGCCCCAAGGACTTGGACCCCGAGCTTTGCACCCACCTCACTACGCTTCGCTGGCATG 209
QY 41 ThrAsnHisGlnLeuSerThrThrGluTrpAsnAspGluThrLeuTrpGlnPheAsn 60
Db 210 ACCAACCCACAGCTGAGCACCACCTGAGTGAATGACGAGACTCTCTACGAGGAGTTCAAT 269
QY 61 GlyLeuLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPhe 80
Db 270 GGCCTGAAGAGATGAATCCCAAGCTGAAGACCTTTAGCCATCGAGGCTGGAAATTC 329
QY 81 GlyThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsn 100
Db 330 AGCACTCAGAGATTCACAGATATGTTAGCCAGGCCCAACACCGTCAGACCTTTGTCAAC 389
QY 101 SerAlaIleArgPheLeuArgLysThrSerPheAspGlyLeuAspLeuAspTrpGluTrp 120
Db 390 TCGGCGCATCAGGTTTCTCGCAAAATACAGCTTTGACGGCTTGACCTTGTGAGTGGAGTAC 449
QY 121 ProGlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAsp 140
Db 450 CCAGGAAGCCAGGGAGGCCCTGCGGTAGACAGGAGCGCTTCACAACTGTTGATCAGGAC 509
QY 141 LeuAlaAsnAlaPheGlnGlnGlnAlaGlnThrSerGlyLysGluArgLeuLeuLeuSer 160
Db 510 TTGGCCCATGCTTCCAGCAGGAGGCCAGACCTCAGGGAAGGACGCTTCTTCTGAGT 569
QY 161 AlaAlaValProAlaGlyGlnThrThrValAspAlaGlyTrpGluValAspLysIleAla 180

Db 570 GCAGCGGTTCCAGCTGGGAGACCTATGTGGATGCTGGATACAGAGTGAGCAAAATCGCC 629
QY 181 GlnAsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysVal 200
Db 630 CAGAACTGGATTGTGCAACCTTATGGCTTACGACTTCCATGGCTCTTGGGAGAGGTC 689
QY 201 ThrGlyHisAsnSerProLeuTyrLysArgGlnGluSerGlyValAlaAlaSerLeu 220
Db 690 ACGGCACATAACAGACCCCTCTACAGAGGCAAGAGAGAGTGGTGACAGCAGCCGCTC 749
QY 221 AsnValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeu 240
Db 750 AACGTGGATGCTGTGCAACAGTGGCTGCAAGAGGAGGAGCCCTCCAGCAAGCTGATC 809
QY 241 LeuGlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerAspThrArgVal 260
Db 810 CTTGGCATGCTTACCTACGAGCGCTCTTCACACTGGCCCTCTCATCAGACACAGAGTG 869
QY 261 GlyAlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyGlyMetLeu 280
Db 870 GGGGCCCCAGCCAGGAGTCTGGCACTCCAGGCCCTTCACCAAGGAGGAGGAGTGGTG 929
QY 281 AlaTyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLys 300
Db 930 GCCTACTATGAAGTCTGCTCTCTGGAAGGGGGGCCACCAACAGAGATCCAGGATCAGA 989
QY 301 ValProTyrIlePheArgAspAsnGlnTrpValGlyPheAspValGluSerPheLys 320
Db 990 GTGCCCTACATCTCCGGCAACACAGTGGGTGGGCTTTGATGATGGAGAGCTTCAAA 1049
QY 321 ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyGlyAlaMetValTrpAlaLeuAsp 340
Db 1050 ACCAAGGTCAGCTATCTGAAGCAGAGGAGCTGGGCGGGGCCATGGTCTGGGCACTGGAC 1109
QY 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeu 360
Db 1110 TTAGATGACTTTGGCGGCTCTCTCTGCAACAGGCGCGATACCCCTCATCCAGACGCTA 1169
QY 361 ArgGlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLys 380
Db 1170 CGCAGGAACTGAGTCTTCATCTTGGCTTCCAGGACCCCGAGAGCTTGAAGTTCACAAA 1229
QY 381 ProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGln 400
Db 1230 CCAGGTTCAGCCCTCTGAACCTGAGCATGGCCCCAGCCCTGGACAGACAGCTTCTGCCAC 1289
QY 401 GlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyrSerCysAla 420
Db 1290 GGCAGAGCTGATGGGCTCTATCCCAATCTCCGGGAACGGTCCAGCTTCTACAGCTGTGCA 1349
QY 421 AlaGlyArgLeuPheGlnSerCysProThrGlyLeuValPheSerAsnSerCysLys 440
Db 1350 GCGGGCGGCTGTTCAGAAAGCTGCCCGACAGGCTGGTGTTCAGCAACTCTTCGCAAA 1409
QY 441 CysCysThrTrpAsn 445
Db 1410 TGCTGCACCTGGAT 1424

RESULT 7

LOCUS AR212139 1656 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 3 from patent US 6399571.
ACCESSION AR212139
VERSION AR212139.1 GI:21515644
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1656)
AUTHORS Gray,P.W. and Tjoelker,L.W.
TITLE Chitinase chitin-binding fragments
JOURNAL Patent: US 6399571-A 3 04-JUN-2002;
FEATURES Location/Qualifiers

source 1. 1656
BASE COUNT 365 a 497 c 447 g 347 t
ORIGIN
Alignment Scores:
Pred. No.: 2,76e-175 Length: 1656
Score: 2392.00 Matches: 444
Percent Similarity: 99.78% Conservative: 0
Best Local Similarity: 99.78% Mismatches: 1
Query Match: 99.75% Indels: 0
DB: 6 Gaps: 0
US-10-004-219B-10 (1-445) x AR212139 (1-1656)
QY 1 AlalysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPhe 20
Db 90 GCAAACTGGTCTGCTACTTCCCAACTGGGCCAGTACACACAGGGGGAGGCTCGCTTC 149
QY 21 LeuProLysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet 40
Db 150 CTGCCAAGGACTTGGACCCCGAGCTTTGGACCCACCTCATCTACGCCCTTGGTGGCATG 209
QY 41 ThrAsnHisGlnLeuSerThrThrGluTrpAsnAspGluThrLeuTyrGlnLupheAsn 60
Db 210 ACCAACCAACAGCTGAGCACCACTGAGTGGAATGACGAGACTCTCTACCAAGAGTTCAAT 269
QY 61 GlyLeuLysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPhe 80
Db 270 GGCCTGAAGAAGATGAAATCCCAAGCTGAAGACCTCTTTAGCCATCGGAGGCTGGAAATTC 329
QY 81 GlyThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsn 100
Db 330 AGCACTCAGAAGTTCCAGATATGTTAGCCAGGCGCAACACCGCTCAGACCTTTGTTCAC 389
QY 101 SerAlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyr 120
Db 390 TCGGCATCAGGTTTCTCGGCAATATACAGCTTTCAGCGCTTGACCTTTCAGCTGGAGTAC 449
QY 121 ProGlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAsp 140
Db 450 CCAGGAAGCCAGGGAGGAGCCCTGCGGTAGACAAGGAGCGCTTCACAAACCTGGTACAGGAC 509
QY 141 LeuAlaAsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuSer 160
Db 510 TTGCCAATGCTTCCAGCAGAGAGCCCGAGCTCAGGGAAGGAACGGCTTCTTCTGAGT 569
QY 161 AlaAlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAla 180
Db 570 GCAGCGGTTCCAGCTGGGCGACACTATGTGGATGCTGGATACGAGGTGGACAAAATCGCC 629
QY 181 GlnAsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysVal 200
Db 630 CAGAACTGGATTTTGTCAACCTTATGGCTTACGACTTTCATGGCTCTTGGGAGAGGTC 689
QY 201 ThrGlyHisAsnSerProLeuTyrLysArgGlnGluSerGlyValAlaAlaSerLeu 220
Db 690 ACGGACATACAGCCCTCTACAGAGGCAAGAGAGTGGTGACAGCAGCCAGCTTC 749
QY 221 AsnValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeu 240
Db 750 AACGTGGATGCTGTGCAACAGTGGCTGCGAGAGGAGGAGCCCTCCAGCAAGCTGATC 809
QY 241 LeuGlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerAspThrArgVal 260
Db 810 CTTGGCATGCTTACCTACGAGCGCTCTTCACACTGGCCCTCTCATCAGACACAGAGTG 869
QY 261 GlyAlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyGlyMetLeu 280
Db 870 GGGGCCCCAGCCAGGAGTCTGGCACTCCAGGCCCTTCACCAAGGAGGAGGAGTGGTG 929
QY 281 AlaTyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLys 300

```
Db 930 GCCTACTGATGAGTCTGCTCTCGAAGGGGGCCACCAACAGAGAAATCCAGGATCAGAG 989
Qy 301 ValProTyrIlePheArgAspAsnGlnTrpValGlyPheAspValGluSerPheLys 320
Db 990 GTCCCTTACATCTTCGGGACACCAAGTGGTGGCTTTGATGATGTGAGAGCTTCAAA 1049
Qy 321 ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyGlyAlaMetValTrpAlaLeuAsp 340
Db 1050 ACCAAGTCTAGCTATCTGAAGCAGAGAGGACTGGCGGGGCCGATGCTTGGGCACCTGGAC 1109
Qy 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeu 360
Db 1110 TTAGATGACTTTCGGGCTCTCTCTGCAACAGAGGGCGGATACCCCTCATCCAGAGCTA 1169
Qy 361 ArgGlnGluLeuSerLeuProTyrLeuProSerGlyThrProGlnLeuGluValProLys 380
Db 1170 CGCAGGAATGAGTCTTCTCATACTTGGCTTTCAGGACCCCGAGGCTTGAAGTTCCAAA 1229
Qy 381 ProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGln 400
Db 1230 CCAGGTGAGCCCTCTGAACCTGAGCATGGCCCCAGCCCTGGACAAGACAGCTTCTGCCAG 1289
Qy 401 GlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyrSerCysAla 420
Db 1290 GGCAGAGCTGATGGCTCTATCCCAATCTCTCGGGAACGGTCCAGCTTCTACAGCTGCA 1349
Qy 421 AlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerAsnSerCysLys 440
Db 1350 GCGGGGGGCTGTTCAGCAAGCTGCCCGACAGGCTGGTGTTCAGCAACTCTCTGCCAAA 1409
Qy 441 CysCysThrTrpAsn 445
Db 1410 TGCTGACCTTGGAAAT 1424

RESULT 8
LOCUS AX108752 1656 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 3 from Patent WO0123430.
ACCESSION AX108752
VERSION AX108752.1 GI:13923944
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1656)
AUTHORS Allison,D.S., Diesch,G.N., Gray,P.W., Shaw,K.D. and Steiner,B.H.
TITLE Human chitinase immunoglobulin fusion proteins
JOURNAL Patent: WO 0123430-A 3 05-APR-2001;
ICOS CORPORATION (US)
FEATURES
source 1..1656
location Location/Qualifiers
/db_xref="taxon:9606"
/organism="Homo sapiens"
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAC37768.1"
/db_xref="GI:13923945"
translation="MVRSAWAGFWVLLMPGSAKLVICYFTNNAQYQGEARFLPK
DLDPSTLHLYAFAGMTNHLSTTWNDETLQEFNGLKNPKLKTLLAIGWNFS
TQKFTDMVATANNROTFFVNSAIRFLKYSFGDLIDWEYFGSQSGSPVDERFTTLVQ
DLANAFQQAQTSQKRLLSAIVPAGQTVFVAGYVDKIAQNLDFVNLMAVDFHGSW
EKVTGNSPLYKQESGGAASLVNDAAVQWLQKCTPASKLILGMPTVGRSFTLASS
TRVUCAPATGSGTGPFFKEGGLAYIEVCWSKQATKQRIQDKVPYIFRDNQWGF
DDVESFTKVSYLKGLGGMWALDDDFAGFCNQGRIPLIQLROELSLPYLPS
GPELVPKPGQSEPEHGPSPQDFTFCQKADGLYPNRPRRSFSSCAAGRLFOQSC
PTGLVFSNSCKCTWN"
mat_peptide 90..1424
/product="unnamed"
BASE COUNT 365 a 497 c 447 g 347 t
ORIGIN
```

```
Alignment Scores: 2.76e-175 Length: 1656
Pred. No.: 2392.00 Matches: 444
Score: 2392.00 Matches: 444
Percent Similarity: 99.78% Conservative: 0
Best Local Similarity: 99.78% Mismatches: 1
Query Match: 99.75% Indels: 0
DB: 6 Gaps: 0
US-10-004-219B-10 (1-445) x AX108752 (1-1656)
Qy 1 AlaLysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPhe 20
Db 90 GCAAACTCTGCTGCTACTCTTCCAACTGGGCCAGTACAGACAGGGGGAGGCTCGCTTC 149
Qy 21 LeuProLysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet 40
Db 150 CTGCCCAAGGACTTGGACCCCGAGCCCTTGGACCCCACTCTCTAGCCCTTCGCTGGCATG 209
Qy 41 ThrAsnHisGlnLeuSerThrThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsn 60
Db 210 ACCAACCCAGCTGAGCACCACTGAGTGGAAATGACGAGACTCTCTACCCAGGAGTTCAAT 269
Qy 61 GlyLeuLysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPhe 80
Db 270 GGCCTGAAGAAGATGAATCCCAAGCTGAAGACCCCTGTAGCCATCGAGGCTGGAAATTC 329
Qy 81 GlyThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsn 100
Db 330 AGCATCTCAGAGTTTACAGATATGTTAGCCAGGCGCAACACCGTCAGACCTTTGTCAAC 389
Qy 101 SerAlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyr 120
Db 390 TCGGCCATCAGGTTTCTCGCAATACAGCTTGTACGGCTTGACCTTGACCTTGAGTGAGTAC 449
Qy 121 ProGlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAsp 140
Db 450 CCAGGAAGCCAGGGAGGCGCTGCGCTAGACAGGAGCGCTTCAACACCCCTGTGTAGAGAC 509
Qy 141 LeuAlaAsnAlaPheGlnGlnGlnAlaGlnThrSerGlyLysGluArgLeuLeuLeuSer 160
Db 510 TTGGCCATGCTTTCAGCAGGAGGAGCCAGACCTCAGGGAGGAGGAGGCTTCTTCTGAGT 569
Qy 161 AlaAlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAla 180
Db 570 GCAGCGGTTCCAGCTGGGCGAGACCTATGTGGATGTGGATACAGGTGGACAAAATCGCC 629
Qy 181 GlnAsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysVal 200
Db 630 CAGAACCTGGATTTGTCAACCTTATGGCTTACGACTTCCATGGCTCTTGGGAGAGGTC 689
Qy 201 ThrGlyHisAsnSerProLeuTyrLysArgGlnGluSerGlyAlaAlaAlaSerLeu 220
Db 690 ACGGACATACAGCCCTCTACAGAGGCGCAAGAGAGAGTGGTGCAGACCCAGCCTC 749
Qy 221 AsnValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuIle 240
Db 750 AACGTGGATGCTGCTGTCACAGTGGCTGCAGAGGGGAGCCCTCCAGCAGCAAGCTGATC 809
Qy 241 LeuGlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgVal 260
Db 810 CTTGGCATGCTTACCTACGAGCGCTCTTCACTACCTGGCTCTCTATCAGACACCAAGAGT 869
Qy 261 GlyAlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyGlyMetLeu 280
Db 870 GGGGGCCAGCAGCAGGCTTGGCACTCCAGGCCCTTCCACCAAGAGGAGGAGGATGCTG 929
Qy 281 AlaTyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLys 300
Db 930 GCCTACTATGAATCTGCTCTCTGGAAGGGGGCCACCAACAGAGAAATCCAGGATCAGAAG 989
Qy 301 ValProTyrIlePheArgAspAsnGlnTrpValGlyPheAspValGluSerPheLys 320
Db 990 GTCCCTTACATCTTCGGGACACCAAGTGGTGGCTTTGATGATGTGAGAGCTTCAAA 1049
```

```

QY 321 ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyGlyAlaMetValTyrAlaLeuAsp 340
DB 1050 ACCAAGGTCAGTATCTGAGCAGAGGAGCTGGCGGGCCATGGTCTGGGCATCTGGAC 1109
QY 341 LeuAspAspPheAlaGlyPheSerCysAlaGlnGlyArgTyrProLeuLeuGlnThrLeu 360
DB 1110 TTAGATGACTTTGGCGGCTCTCTCTGCAACACGAGGCGCATACCCCTCATCCAGACGCTA 1169
QY 361 ArgGlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLys 380
DB 1170 CCGCAGGAAGTCTGAGTCTTCACTACTTCTTCCAGGACCCCGAGCTTGAAGTTCACAAA 1229
QY 381 ProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGln 400
DB 1230 CCAGGTGAGCCCTGAACTGAGCATGCGCCCGCCCTGGACAAGACAGCTTCTGCCAG 1289
QY 401 GlyLysAlaAspGlyLeuTyrProLysProLysProLysProLysProLysProLys 420
DB 1290 GGCAGGCTGATGGGCTCTATCCCAATCTCGGGAACGGTCCAGCTTCTACAGCTGTGCA 1349
QY 421 AlaGlyArgLeuPheGlnSerCysProThrGlyLeuValPheSerAsnSerCysLys 440
DB 1350 GCGGGCGGCTGCTTCCAGAAAGCTGCCCGACAGCGCTGTGTTCAGCAACTCTTGCAAA 1409
QY 441 CysCysThrTyrAsn 445
DB 1410 TGCTGCACCTGGAT 1424

RESULT 9
HSU62662 1710 bp mRNA linear PRI 07-SEP-2000
LOCUS Homo sapiens chitotriosidase precursor, mRNA, complete cds,
DEFINITION U62662
ACCESSION U62662
VERSION U62662.1 GI:9988475
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1710)
AUTHORS Boot,R.G., Renkema,G.H., Strijland,A., van Zonneveld,A.J. and Aerts,J.M.
TITLE Cloning of a cDNA encoding chitotriosidase, a human chitinase produced by macrophages
J. Biol. Chem. 270 (44), 26252-26256 (1995)
MEDLINE 96064695
PUBMED 7592832
REFERENCE 2 (bases 1 to 1710)
AUTHORS Boot,R.G., Renkema,G.H., Strijland,A. and Aerts,J.M.F.G.
TITLE Direct Submission
JOURNAL Submitted (01-JUL-1996) Department of Biochemistry, AMC, University of Amsterdam, Meibergdreef 15, Amsterdam 1105 AZ, The Netherlands
REFERENCE 3 (bases 1 to 1710)
AUTHORS Boot,R.G., Renkema,G.H., Verhoeck,M., Strijland,A., Bliek,J., de Meulemeester,T.M., Mannens,M.M. and Aerts,J.M.
TITLE The human chitotriosidase gene. Nature of inherited enzyme deficiency
J. Biol. Chem. 273 (40), 25680-25685 (1998)
MEDLINE 98421482
PUBMED 9748235
FEATURES
source
1..1710
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Ch3.1"
/cell_type="macrophage"
/clone_lib="macrophage-1 cDNA library of Rolf G. Boot"
10..1173
note="splice variant 40 kDa"
/codon_start=1
/product="chitotriosidase precursor"

```

```

/protein_id="AAG10644.1"
/db_xref="GI:9988476"
/translation="MVRSVAMGFMVLLMIWMSAAKLVCYFTWAOYRQGEARFLPK
DLDSLCTHLIYAFAGNTHOLSTEWDEFLYQEFNGLKMNPKLKTLLAIGWNFG
TQKFTDWATANNQITFVNSAIRFLRKYSFDGLDWEYDPSQSGSPAVDKERFTLLVQ
DLNAPQEAQTSKGRLLLSAAVPAQGTYYDAGTEVDKIAONLDFVNLMAIDFHSW
SKVTGNSPLYRQEESSAASLNDAVQOQWLOKGTPTASKLILGMPYTGSRFTLASS
EDTVRGAPATGSGTPTGPTKEGMLAYVEVCWKGATKQRIQDQKVPYIFRDNOWVG
DDVESFTKSVYLAQKGLGGAMWALDLDPDGFSQNGRYPLIQTLELNG"
10..72
sig_peptide
mat_peptide 73..1170
misc_feature 1166..1236
/feature="chitotriosidase"
/feature="note=alternatively spliced exon, absent in 50 kDa
chitotriosidase (GenBank Accession Number U29615),
fragment contains stop codon, therefore coding for a
protein of 40 kDa"
polyA_signal 1673..1678
polyA_site 1698..1710
BASE COUNT 379 a 464 g 364 t
ORIGIN
Alignment Scores:
Pred. No.: 3.19e-172 Length: 1710
Score: 2352.50 Matches: 444
Percent Similarity: 94.67% Conservative: 0
Best Local Similarity: 94.67% Mismatches: 1
Query Match: 98.10% Indels: 25
DB: 9 Gaps: 1
US-10-004-219B-10 (1-445) x HSU62662 (1-1710)
QY 1 AlaLysLeuValCysTyrPheThrAsnTyrAlaGlnTyrArgGlnGlyAlaArgPhe 20
DB 73 GCAAACTGGTCTGCTACTTCCACAACTGGGCCCGAGTACAGACGGGAGGCTCGCTTC 132
QY 21 LeuProLysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet 40
DB 133 CTGCCCAAGGACTTGGACCCCGAGCCCTTTGCACCCACCTCATCTACGCTTCGCTGGCATG 192
QY 41 ThrAsnHisGlnLeuSerThrThrGluTyrAsnAspGluThrLeuTyrGlnGluPheAsn 60
DB 193 ACCAACCAACAGCTGAGCACCACCTGAGTGGATGAGAGACTCTCTACCGAGAGTTCAAT 252
QY 61 GlyLeuLysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTyrAsnPhe 80
DB 253 GGCCTGAAGAAGATGAATCCCAAGCTGAAGACCCCTGTAGCCATCGAGGCTGGAAATTC 312
QY 81 GlyThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsn 100
DB 313 GGCACCTCAGAAGTTCCACAGATATGCTAGCCAGCGCCCAACACCGTCAGACCTTTGTCAC 372
QY 101 SerAlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTyrGluTyr 120
DB 373 TCGGCATCAGGTTTCTCGCAAAATACAGCTTGTGACGCCCTTGACCTTGACCTGGAGTAC 432
QY 121 ProGlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAsp 140
DB 433 CCAGGAACCGAGGGAGGCCCTGCGGTAGACAGGAGCCCTTCACAAACCTCGTGTGACGAC 492
QY 141 LeuAlaAsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuSer 160
DB 493 TTGGCCAAATGCTTCCAGCAGGAGGCCAGACCTCAGGAGGAACGCTTCTCTCTGAGT 552
QY 161 AlaAlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAla 180
DB 553 GCAGCGGTTCAGCTGGGACAGACCTATGTGGATGTGTGATACGAGGTGACAAAATCGCC 612
QY 181 GlnAsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTyrGluLysVal 200
DB 613 CAGAACCTGGGATTTGTCAACCTATATGCGCTACGACTTCCATGGCTCTTGGGAGAGGTC 672
QY 201 ThrGlyHisAsnSerProLeuTyrLysArgGlnGluGluSerGlyAlaAlaAlaSerLeu 220

```

Db 673 ACGGACATAACAGCCCTCTACAAGAGCGCAAGAGAGAGTGGTGCAGCAGCCAGCTTC 732
Qy 221 AsnValaspAlaalaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuLeu 240
Db 733 AAGTGGATGCTGCTGTGAACAGTGGCTGCAAGAGGGACCCCTGCCAGCAGCTGATC 792
Qy 241 LeuGlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgVal 260
Db 793 CTGGGATGCTACCTACGAGCGCTCTTCCACACTGGCCCTCTCATCACACACAGAGTG 852
Qy 261 GlyAlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyMetLeu 280
Db 853 GGGGCCCCAGCCAGGGCTGGCACTCCAGGGCCCTTCAACAGGAAGAGGGATGCTG 912
Qy 281 AlaTyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgLysGlnAspGlnLys 300
Db 913 GCTACTATGAAGTCTGCTCTCTGGAAAGGGGGCCACCAACAGAGAATCCAGGATCAGAAG 972
Qy 301 ValProTyrIlePheArgAspAsnGlnTrpValGlyPheAspValGluSerPheLys 320
Db 973 GTGCCCTACATCTTCGGGACACACAGTGGTGGCTTTGATGATGTGGAGAGCTTCAAA 1032
Qy 321 ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyGlyAlaMetValTrpAlaLeuAsp 340
Db 1033 ACCAAGGTGAGCTATCTGAAGCAGAGAGGAGTGGGGGGCCCATGCTCTGGCACTGGAC 1092
Qy 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuLeuGlnThrLeu 360
Db 1093 TTAGATGACTTTGGCCGGCTTCTCCTGCACACAGGGCCGATACCCCTCATCCAGAGCTA 1152
Qy 361 ArgGlnGluLeuSer----- 365
Db 1153 CGGCGGAAGCTGAA-TGGGTAAAGCCTTAACCTGCTGCATGTGAGGCCAGGTGTTCC 1211
Qy 366 -----LeuProTyrLeuProSerGlyThrProGluLeu 376
Db 1212 CTGTGGCACTGTCTTTCAGCTGTAGGCTTTCATATCTTCCATCTTTCAGGCACCCAGAGCTT 1271
Qy 377 GluValProLysProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAsp 396
Db 1272 GAAGTTCCTCAAAACAGGTGAGCTGAGCTCTGAACTGAGCATGGCCCCAGCCCTGGACAAGAC 1331
Qy 397 ThrPheCysGlnGlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerPhe 416
Db 1332 ACCTTCTGCCAGGGCAAGCTGATGGCTCTATCCCAATCTCTGGGAACGGTCCAGCTTC 1391
Qy 417 TyrSerCysAlaAlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSer 436
Db 1392 TACAGCTGTGCAGCGGGGGCTGTTCCAGCAAAAGCTGCCCAAGCTGCCAGAGCCCTGGTTCAGC 1451
Qy 437 AsnSerCysLysCysCysThrTrpAsn 445
Db 1452 AACTCTCTGCAATGCTGCACCTGGAAT 1478

RESULT 10

LOCUS AR172108 1713 bp DNA linear PAT 17-DEC-2001
DEFINITION Sequence 5 from patent US 6303118.
ACCESSION AR172108
VERSION AR172108.1 GI:17911599
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1713)
Aerts, J. Maria, Francisus, Gerardus.
Human chitinase, its recombinant production, its use for
decomposing chitin, its use in therapy or prophylaxis against
infection diseases
JOURNAL Patent: US 6303118-A 5 16-OCT-2001;
FEATURES Location/Qualifiers
source 1. .1713

BASE COUNT 380 a 504 c 464 g 365 t
ORIGIN
Alignment Scores:
Score: 6.49e-172 Length: 1713
Pred. No.: 2348.50 Matches: 443
Percent Similarity: 94.66% Conservative: 0
Best Local Similarity: 94.66% Mismatches: 1
Query Match: 97.94% Indels: 25
DB: 6 Gaps: 1

US-10-004-219B-10 (1-445) x AR172108 (1-1713)

Qy 2 LysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPheLeu 21
Db 79 AAACCTGGTCTGCTACTTCCACCACTGGGCCCNAGTACACAGAGGGGGAGGCTCGCTTCCTG 138
Qy 22 ProLysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMetThr 41
Db 139 CCCAAGGACTTGGACCCCGAGCCTTTGCACCCACCTCATCTACGCCCTTCGCTGGCATGACC 198
Qy 42 AsnHisGlnLeuSerThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsnGly 61
Db 199 AACCCAGCTGAGCACCACCTGAGTGGAAATGACGAGACTCTCTACAGAGATCAATGGC 258
Qy 62 LeuLysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTyrAsnPheGly 81
Db 259 CTGAAGAGATGAATCCCAAGCTGNAAGCCTGTGTAGCCATCGGAGGCTGGATTTTCGGC 318
Qy 82 ThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsnSer 101
Db 319 ACTCAGAAGTTTCACAGATATGTTAGTACCGCCCAACCAACCCGTACAGACCTTTGTCAACTCG 378
Qy 102 AlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyrPro 121
Db 379 GCCATCAGGTTTCTCGCAAAATACAGCTTTCAGGCCCTTGACCTTGACTGGGAGTACCCA 438
Qy 122 GlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAspLeu 141
Db 439 GGAAGCCAGGGGAGCCCTGCCGTAGACAAGGAGCGCTTCAACAACCTTGGTACAGGACTTG 498
Qy 142 AlaAsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuSerAla 161
Db 499 GCAATGCTCTCCAGCAGGAAGCCAGACCTCAGGGAAGGAACGCTTCTCTCTGAGTGCA 558
Qy 162 AlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAlaGln 181
Db 559 GCGGTTCAGCTGGGCAGACCTATGTGGATACGAGGTGGACAAATATGCCAG 618
Qy 182 AsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysValThr 201
Db 619 AACCTGGATTTTGTCAACCTTATGGCTTACGACTTCATGGCTCTTGGAGAGGTACAG 678
Qy 202 GlyHisAsnSerProLeuTyrLysArgGlnGluLysSerGlyAlaAlaSerLeuAsn 221
Db 679 GGACATAACAGCCCTCTACAAGAGCAAGAGAGTGGTGCAGCAGCCAGCCCTCAAC 738
Qy 222 ValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuLeu 241
Db 739 GTGGATGCTGCTGTGCAACAGTGGCTGCAGAGGGGACCCCTGCCAGAGCTGATCTT 798
Qy 242 GlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgValGly 261
Db 799 GCGATGCCCTACTAGGAGCGCTCTTCACTGGCTCTCTCATCAGACACAGAGGTGGG 858
Qy 262 AlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyMetLeuAla 281
Db 859 GCGCCAGCCACAGGGTCTGGCACTCCAGGCCCTTTCACCAAGGAAGAGGAGGTGCTGCC 918
Qy 282 TyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLysVal 301
Db 919 TACTATGAAGTCTGCTCTCTGGAAGGGGGCCCAACAGAGATCCAGATCAGAGGTG 978

HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: RAB and HRI.

FEATURES
source Location/Qualifiers
1..1599

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CD34C2000051"
/cell_type="CD34+ Cells"
/clone_lib="CD34C2"
/note="Cloning vector: pME18SFL3-primary culture, CD34+ Cells"

BASE COUNT 339 a 481 c 437 g 342 t
ORIGIN

Alignment Scores:

Pred. No.: 6,12e-166 Length: 1599
Score: 2270.50 Matches: 425
Percent Similarity: 95.51% Conservat: 0
Best Local Similarity: 95.51% Mismatches: 1
Query Match: 94.68% Indels: 19
DB: 9 Gaps: 1

US-10-004-219B-10 (1-445) x AK055165 (1-1599)

QY 1 AlaLysLeuValCysThrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPhe 20
DB 100 GC AAAACTGGTCTGCTACTTCAACCACTGGGCCAGTACACAGCGGGAGGCTCGCTTC 159
QY 21 LeuProLysAspLeuAspProSerLeuCysThrHisLeuLeuTyrAlaPheAlaGlyMet 40
DB 160 CTGCCCAAGGAGCTTGGAGCCGCCACCTTTGCACCCACCTCATCTACGCCCTTCGCTGCATG 219
QY 41 ThrAsnHisGlnLeuSerThrThrClnTrpAsnAspGluThrLeuTyrClnGluPheAsn 60
DB 220 ACCAACCAACAGCTGAGCACCACCTGAGTGGATACGAGACTCTTACAGAGATTCAAT 279
QY 61 GlyLeuLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPhe 80
DB 280 GGCCTGAAGAGATG----- 294
QY 81 GlyThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsn 100
DB 295 -----TTCACAGATATGTTAGCCAGCGCCAAACCAACCGTCAGACCTTTGTCAAC 342
QY 101 SerAlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyr 120
DB 343 TCGGCCATCAGGTTTCTGCGCAATATACAGCTTTGACGCCCTTGACTTGGAGGTAC 402
QY 121 ProGlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAsp 140
DB 403 CCAGGAAGCCAGGGAGGCCCTGCCGTAGCAAGGAGCGCTTCACAACTTGGTGCAGGAC 462
QY 141 LeuAlaAsnAlaPheGlnGlnGlnAlaGlnThrSerGlyLysGluArgLeuLeuSer 160
DB 463 TTGGCCAATGCTTCCAGCAGGAAGCCAGACCTCAGGGAAGAAACGCTTCTTCTGAGT 522
QY 161 AlaAlaValProAlaGlyGlnThrThrValAspAlaGlyTyrGluValAspLysIleAla 180
DB 523 GCAGCGGTTCAGCTGGGCGACACCTATGTGATGTGTGATGAGGTGAGGACAAAATCGCC 582
QY 181 GlnAsnLeuAspPheValAsnLeuMetAlaTrpAspPheHisGlySerTrpGluLysVal 200
DB 583 CAGAACCTGGATTTGTCAACCTTATGGCCCTACGACTTCCATGGCTCTTGGGAGAGGTC 642
QY 201 ThrGlyHisAsnSerProLeuTyrLysArgGlnGluSerGlyValAlaAlaSerLeu 220
DB 643 ACGGACATAACAGCCCTCTTACAAGAGGCAAGAGAGTGGTGCAGCAGCCAGCCTC 702
QY 221 AsnValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeu 240
DB 703 AACGTGGATGCTGTGTCACAGTGGCTGCAGAGAGGGGCCCTTCCAGCAGCAAGTATC 762

QY 302 ProTyrIlePheArgAspAsnGlnTrpValGlyPheAspAspValGluSerPheLysThr 321
DB 979 CCTCATCTTCGGGACACACAGTGGTGGCTTTGATGATGGAGAGCTTCAAAACC 1038
QY 322 LysValSerTyrLeuLysGlnLysGlyLeuGlyAlaMetValTrpAlaLeuAspLeu 341
DB 1039 AAGGTGAGTATCTGAAGCAGAGGAGACTGGCGGGGCCCATGCTGTGGCACTGGACTTA 1098
QY 342 AspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeuArg 361
DB 1099 GATGACTTTGGCGGCTTCTCTGCAACAGGCGCCGATACCCCTCATCAGAGCTACGG 1158
QY 362 GlnGluLeuSer----- 365
DB 1159 CAGGAAGTAAAGCTTAACTGCCTGTGCATGTGAGGCCAGGTGTTGCTCG 1217
QY 366 -----LeuProTyrLeuProSerGlyThrProGluLeuGlu 377
DB 1218 TGGCACTGTGCTTACAGTGTAGTCTTCCATACTTGCCTTTCAGGCACCCAGAGCTTGA 1277
QY 378 ValProLysProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThr 397
DB 1278 GTTCAAAACCAAGGTCAGCCCTCTGAACCTGAGCATGGCCCAAGCCCTGGCAAGACAG 1337
QY 398 PheCysGlnGlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyr 417
DB 1338 TTCTGCCAGGCAAGCTGATGGCTCTATCCATCTCGGGAAGCTTCCAGCTTCTAC 1397
QY 418 SerCysAlaAlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerAsn 437
DB 1398 AGCTGTGCAGCGGGCGGCTGTTCAGCAAAAGCTGCCGACAGGCGTGGTGTTCAGCAAC 1457
QY 438 SerCysLysCysCysThrTrpAsn 445
DB 1458 TCCTGCAATGCTGCACCTGGAAT 1481

RESULT 11
LOCUS AK055165
DEFINITION Homo sapiens cDNA FLJ30603 fis, clone CD34C2000051, highly similar to Human chitotriosidase precursor mRNA.
ACCESSION AK055165
VERSION AK055165.1 GI:16549834
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens CD34+ Cells cDNA to mRNA, clone_lib:CD34C2
clone:CD34C2000051.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Ninomiya,K., Wagatsuma,M., Kanda,K., Kondo,H., Yokoi,T., Kodaira,H., Furuya,T., Takahashi,M., Kikkawa,E., Omura,Y., Abe,K., Kamihara,K., Katsuta,N., Sato,K., Tanikawa,M., Yamazaki,M., Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Murakawa,K., Kanehori,K., Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S., Nagahari,K., Masuno,Y., Nagai,K. and Isogai,T.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 1599)
Isogai,T., Otsuki,T. and Sugiyama,T.
Direct Submission
Submitted (24-Oct-2001) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,


```
QY 241 LeuGlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgVal 260
Db 763 CTTGGGATGCTTACCTACGGAGCTCTTACACAGTGGCTCTCATCAGACACAGAGTG 822
QY 261 GlyAlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyMetLeu 280
Db 823 GGGGCCCCAGCCAGCGGTGTGGCACTCCAGGCCCCCTTACCAGAGGAGGAGGATGCTG 882
QY 281 AlaTyrTyrGluValCysSerTyrLysGlyAlaThrLysGlnArgLysGlnAspGlnLys 300
Db 883 GCGTACTATGAAGTCTGTCTCTGGAAGGGGGCCACCAACAGAGATTCAGGATCAGAAG 942
QY 301 ValProTyrIlePheArgAspAsnGlnTyrValGlyPheAspAspValGluSerPheLys 320
Db 943 GTGCCCTACATCTTCGGGACACACAGTGGGTGGGCTTTGATGATGTGGAGAGCTTCAA 1002
QY 321 ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyGlyAlaMetValTrpAlaLeuAsp 340
Db 1003 ACCAAGGTGAGCTATCTGAAGCAGAGAGGAGTGGGGGGGCCCATGTGCTGGGCACCTGGAC 1062
QY 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeu 360
Db 1063 TTAGATGAGCTTTCGGGCTCTCTCTGCAACAGGGCCGATACCCCCCTCATCCAGAGCTA 1122
QY 361 ArgGlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLys 380
Db 1123 CGGCGAGACTGAGTCTTCATACTTTCCTTCAGGCACCCAGAGCTTGAAGTTCCAAA 1182
QY 381 ProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGln 400
Db 1183 CCAGGTGAGCCCTCTCAACCTGAGCATGGCCCGCCCTGGACAGACACGCTTCTGCCAG 1242
QY 401 GlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerPheTyrSerCysAla 420
Db 1243 GGCAAGCTGATGGGCTCTATCCATCTCCGGAACGGTCCAGCTTCTACAGCTGTGCA 1302
QY 421 AlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerAsnSerCysLys 440
Db 1303 GGGGGGGGCTGTTCAGCAAAAGCTGCCGACAGGCTGTGTTCAGCAACTCTCTCAA 1362
QY 441 CysCysThrTrpAsn 445
Db 1363 TGCTGCACCTGGAAT 1377

RESULT 12
AF290004
LOCUS
DEFINITION Homo sapiens acidic mammalian chitinase precursor, mRNA, complete cds.
ACCESSION AF290004
VERSION AF290004.1 GI:12597282
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Boot,R.G., Blommaert,E.F., Swart,E., Ghauharali-van der Vlugt,K.,
Bi,jl,N., Moe,C., Place,A. and Aerts,J.M.
TITLE Identification of a novel acidic mammalian chitinase distinct from
chitotriosidase
J. Biol. Chem. 276 (9), 6770-6778 (2001)
JOURNAL
MEDLINE 21125893
PUBMED 11085997
REFERENCE
AUTHORS Boot,R.G., Verhoeck,M., Swart,E. and Aerts,J.M.
TITLE Direct Submission
JOURNAL Submitted (26-JUL-2000) Dept. of Biochemistry, Academic Medical
Center, University of Amsterdam, Meibergdreef 15, Amsterdam 1105
AZ, The Netherlands
FEATURES
source Location/Qualifiers
1..1625
/organism="Homo sapiens"
```

```
/db_xref="taxon:9606"
/tissue_type="stomach; lung"
104..1534
/EC_number="3.2.1.14"
/Note="chitinase precursor; AMCase; chitinase activity
demonstrated by expression in COS cells; similar to
sequences presented in GenBank Accession Numbers AB025008
and AB025009"
/codon_start=1
/product="acidic mammalian chitinase precursor"
/protein_id="AAG60019.1"
/db_xref="GI:12597282"
/translation="MTKLILLTGLVLLNLQSGAYQLTCYFTNWAQYRPGLRMPD
NIDPLCTHLIYAFAGRONQITTEWNVLYQAFNGLKKNKSLKLLAIGWNFG
TAPFTAMYSTPENQTFITSVIKFLRYQFDFDQDWEYPGSRSPDQKHLFTLVQ
EMRAFEQEAQKQINKPRLMTVAAGAISNQSQVEIPLQSOYLDYIHMTVDLHGSW
EGYTGENSEPLYKYPTDTSNAYLVNVMYKONGAPAEKLIYGFPTYGHNFLSNP
SNTGKSPDGAGPAGPAGKESGIWAYYEICTFLKNGATQWDAPQEVPIAYQGNWV
GYDNKSPDIKAQMLKHNKFGMAWVAILDDFTGFCNQKFPFLITLTKALGQSA
SCTPAQPIETPTTAAPSGSGNGSGSSSSSGSGSGGFCVAVRANGLYPVANRRNAFMC
VNGVTYQNCQAGLVFDTSCDCNWA"
sig_peptide 104..166
BASE COUNT 387 a 468 c 399 g 371 t
ORIGIN
Alignment Scores:
Pred. No.: 6,75e-93 Length: 1625
Score: 1323.00 Matches: 240
Percent Similarity: 67.3% Conservative: 70
Best Local Similarity: 52.1% Mismatches: 126
Query Match: 55.1% Indels: 24
DB: 9 Gaps: 3
US-10-004-219B-10 (1-445) x AF290004 (1-1625)
QY 2 LysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPheLeu 21
Db 170 CAGCTGACATGCTACTTACCACCTGGGCCAGTACCGGCCAGGCTGGGGCGCTTCATG 229
QY 22 ProLysAspLeuAspProSerLeuThrHisLeuIleTyrAlaPheAlaGlyMetThr 41
Db 230 CTGCAACATCGACCCCTGCTCTGTACCCACTGATCTAGCCCTTTGCTGGGAGGAG 289
QY 42 AsnHisGlnLeuSerThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsnGly 61
Db 290 AACACGAGATCACCACTCGATGGAACGATGACTCTCTACCAAGCTTCAATGCG 349
QY 62 LeuLysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnGly 81
Db 350 CTGAAAAATAAGACAGCCAGCTGAAAACTCTCTGGCCATTTGGAGGCTGGAACTTCGGG 409
QY 82 ThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsnSer 101
Db 410 ACTGCCCTTTTCACTGCCATGTTTCTACTCTGAGAACCCGACAGCTTTCATCCTCA 469
QY 102 AlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyrPro 121
Db 470 GTCATCAATTTCTGGCCAGTATGAGTTTGACGGCTGGACTTTTACTGGGAGTACCT 529
QY 122 GlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAspLeu 141
Db 530 GGCTCTCTGGGAGGCCCTCTCAGGACCAAGCATCTCTTCACTGTCTGTGGTGGAGAAATG 589
QY 142 AlaAsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuSerAla 161
Db 590 COTGAAGCTTTTGACAGAGGCGCAAGAGATCAACAAGCCAGGCTGATGCTGCTGCT 649
QY 162 AlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAlaGln 181
Db 650 GCAGTAGCTGTGGCATCTCCAATATCCAGTCTGCTGATGATGATGATGATGATGATGAT 709
QY 182 AsnLeuAppPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysValThr 201
```

CDS

Db	710	TACCTGGACTACATCATGCTCATGACCTACGACCTCCATGGCTCTCTGGGAGGCTACACT	769	FEATURES	Location/Qualifiers	Diego, 9500 Gilman Drive, Mail Code 0368, La Jolla, CA 92093-03687,
QY	202	GlyHisAsnSerProLeuTyrLysArgGlnGluSerGlyValAlaAlaSerLeuAsn	221	source	1..1526	USA
Db	770	GGAGAACAGACCCCTCTACAAATACCGACTGACCCGGACGACCCCTACCTCAAT	829		/organism="Mus musculus"	
QY	222	ValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuLeu	241		/strain="ATCC147668"	
Db	830	GTGGATTATGTCATGACTACTTGAAGGACATGAGCACCAGCTGAGAAGCTCATCGTT	889		/db_xref="ATCC:147668"	
QY	242	GlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerAspThrArgValGly	261		/db_xref="taxon:10090"	
Db	890	GGATTCCCTACCTATGGACACAACTTCATCTCGAGCAACCCCTCCAACACTGGAATGCT	949		/clone="IMAGE 1209227"	
QY	262	AlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyMetLeuAla	281		/tissue_type="skin"	
Db	950	GGCCCCACCTCTGGTCTGCTGCTGGCCCTATGCCACAGAGTCTGGGATCTGGGCT	1009	gene	<1..1526	
QY	282	TyTyrGluValCysSerTrp-----LysGlyAlaThrLysGlnArgIleGlnAspGln	299	CDS	/gene="YNL"	
Db	1010	TACTACGAGATCTGATCTCTCGAAATATGAGCCACTCAGGATGGATGCCCTCAG	1069		<1..1419	
QY	300	LysValProTyrIlePheArgAspGlnTrpValGlyPheAspAspValGluSerPhe	319		/codon_start=1	
Db	1070	GAAGTGCCTTATGCTCATCAGGCAATGTGTGGTTGGCTATGACAAACATCAAGAGCTTC	1129		/product="putative chitinase precursor"	
QY	320	LysThrLysValSerTyrLeuLysGlnLysGlyLeuGlyValAlaMetValTrpAlaLeu	339		/db_xref="GI:6934190"	
Db	1130	GATATTAGGCTCAATGGCTTAAGCAACAATAATTGGAGGCGCCATGCTCGGCCATT	1189	sig_peptide	/translation="AKLLVLTGLALLNAQGSAYNLICYFTNMAQYRPGISFXPDD	
QY	340	AspLeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThr	359	mat_peptide	INPCLTHLIYAFAGMONEITTIEMNDVTLYKAFNDLKNRNSKLKTLAIGGWNFGT	
Db	1190	GATCTGGATGACTTCACTGGCACTTCTGCACACAGGCAAGTTCCCTATATCCACC	1249	misc_feature	APPTTWTSTONROTFTSVIKFLOYGFDGLDLDWEYVPSGRSPDOKHLPTLVLKE	
QY	360	LeuArgGlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValPro	379	misc_feature	MREAFQEQAIESNRPLMTAAGGINSIQAGYEIPELSKYLDPIHVMYDLHGSWE	
Db	1250	CTGAAGAAGGCCCTC-----GGCCTGCAGAGTGCAGTTGCACG	1288		GITGENSLYKIPITETGSNAYLNVDYVNNYNNNGAPAEKLIIVGPEPCHTFLRNPS	
QY	380	LysProGlyGlnProSerGluPro-----	387		DNGIAPSGDGPAGPYTRQAGFWAYEICTFLRSGATEVMDASOEVPYAKANESWL	
Db	1289	GCTCCAGCTCAGCCCATGAGCAATAACTGCTGCCAGTGGCAGCGGGAACGGGAGC	1348		YDNIKSPSVKQWLKQNNFGAMTWAIDDDFTGSCFDQGRFPLTSTLNKALGISTEG	
QY	388	-----GluHisGlyProSerProGlyGlnAspThrPheCysGlnGlyLysAlaAsp	404		CADPVPSEPVTPPGSGSGSGSGSGSGSGGSCGSCGSCGSCGSCGSCGSCGSCGSCG	
Db	1349	GGAGTAGCAGCTCTGAGGAGCAGCTCGGAGCAGCGAGTCTGCTGTCCAGGCCAAC	1408		TYQHQCOAGLVFDYSCNCNP"	
QY	405	GlyLeuTyrProAsnProArgGluArgSerPheTyrSerCysAlaAlaGlyArgLeu	424		<1..60	
Db	1409	GGCCTCTACCCGCTGGCAATAACAGAAATGCTTCTGGCACTGCTGAATGGAGTCACG	1468		/gene="YNL"	
QY	425	PheGlnGlnSerCysProThrGlyLeuValPheSerAsnSerCysLysCysCysThrTrp	444		61..1416	
Db	1469	TACCAGCAAACTGCCAGCGCGGCTTGTCTTCGACACCAGCTGTGATTGCTGCAACTGG	1528		/gene="YNL"	
RESULT 13					394..417	
AF154571					/gene="YNL"	
LOCUS					/note="putative conserved region; active site"	
DEFINITION					415..417	
ACCESSION					/gene="YNL"	
VERSION					/note="putative invariant Glu; active site"	
KEYWORDS						
SOURCE						
ORGANISM						
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						
AUTHORS						
TITLE						
JOURNAL						

Db 367 ATCAAAATTTTCGCTCAGTATGGTTGGATGGACTGGACCTGGAGTACCGGAAATACCCAGGC 426
QY 123 SerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAspLeuAla 142
Db 427 TCACGTGGGAGCCCTCTCAGGACAGCATCTCTTCACTGTCTCTGTGAGGAAATGCCT 486
QY 143 AsnAlaPheGlnGlnGlnAlaGlnThrSerGlyLysGluArgLeuLeuSerAlaAla 162
Db 487 GAAGCTTTTGGCAGGAGGCTATTGAGAGCACAGGCCAGCATGATGGTTACTGTCTGCT 546
QY 163 ValProAlaGlyGlnThrThrValAspAlaGlyTyrGluValAspLysLeuAlaGlnAsn 182
Db 547 GTAGCTGGTGGGATTTCCAACTCCAGGCTGGCTATGAGATCCCTCAACTTTCTAGATAC 606
QY 183 LeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTyrGluLysValThrGly 202
Db 607 CTGGATTTTCATCATGTCATGATGATGATGATGATGATGATGATGATGATGATGATG 666
QY 203 HisAsnSerProLeuTyrLysArgGlnGlnGluSerGlyAlaAlaAlaSerLeuAsnVal 222
Db 667 GAGATAGTCTCTTTACAAATACCTCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCT 726
QY 223 AspAlaAlaValGlnGlnThrLeuGlnLysGlyThrProAlaSerLysLeuLeuLeuGly 242
Db 727 GATTGTGTCATGAACATTTTGAAGAACAAATGGAGCCCTGAGAGCTGAGAGCTGATG 786
QY 243 MetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgValGlyAla 262
Db 787 TTCCAGAGATGAGACACCTTCATCTCTGAGAACCCCTCTGATATGATGATGATGATG 846
QY 263 ProAlaThrGlySerGlyThrProGlyProPheThrLysGlyGlyGlyMetLeuAlaTyr 282
Db 847 CTTACCTCTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 906
QY 283 TyrGluValCysSerTrp-----LysGlyAlaThrLysGlnArgIleGlnAspGlnLys 300
Db 907 TATGAGATTTGACCTTTCTGAGAGTGGAGCCACCTGAGCTGGGATGCTGCCCAAGAA 966
QY 301 ValProTyrIlePheArgAspAsnGlnTyrValGlyPheAspAspValGluSerPheLys 320
Db 967 GTGCCCTATGCTTATAGGCCACAGAGTGGCTGGCTGATGACATATCAAGAGCTTCAGT 1026
QY 321 ThrLysValSerTyrLeuLysGlnLysGlyLysGlyGlyAlaMetValTrpAlaLeuAsp 340
Db 1027 GTTAAGGCTCAGTGGCTTAAGCAGAACAAATTTGGAGGTGCCATGATCTGGGCCATTGAC 1086
QY 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeu 360
Db 1087 CTTGTAGTACTTCACTGGCTCTTCTGTGATCAGGAAATTTCTCTGACTCTCTACTTTG 1146
QY 361 ArgGlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLys 380
Db 1147 AACAAAGCCCTTGGCATATCCACTGAGGTTGACAGCTCTCTGACGTG-----Glu 388
QY 381 ProGlyGlnProSerGluPro-----CCTTCGAGCAGTACTCTCTCAGAGTGGGAGTGAAGC 1245
Db 389 HisGlyProSerProGlyGlnAspThrPheCysGlnGlyAlaAspGlyLeuTyrPro 408
Db 1246 TCCGGAGGAGCTCTGGAGCAGTGAATCTGTGCGCAGACAAAGCAGATGGCTCTACCT 1305
QY 409 AsnProArgGluArgSerSerPheTyrSerCysAlaAlaGlyArgPheGlnGlnSer 428
Db 1306 GTGGCAGATCAGAGAATGCTTTTGGCAGTGCATCAATGGAATCATACACGAGCAT 1365
QY 429 CysProThrGlyLeuValPheSerAsnSerCysLysCysCysThrTrp 444
Db 1366 TGTCAAGCAGGCGCTGTTTTCATACACAGCTGTAATTTGTCGAACGTG 1413

RESULT 14

BC011134

LOCUS

BC011134 1538 bp mRNA linear ROD 07-AUG-2002

DEFINITION
Mus musculus, similar to eosinophil chemotactic cytokine, clone
MGC:18771 IMAGE:4165150, mRNA, complete cds.
ACCESSION
BC011134
VERSION
BC011134.1 GI:15029821
KEYWORDS
MGC.
SOURCE
house mouse.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Strausberg, R.
Direct Submission
Submitted (25-JUL-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabbs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gonaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 24 Row: e Column: 11.

FEATURES
Location/Qualifiers

1..1538
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="FVB/N"
/clone="MGC:18771 IMAGE:4165150"
/tissue_type="Salivary gland, 10 week old female mouse"
/clone_lib="NCI CGAP_SG2"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
323..1420
/codon_start=1
/product="Similar to eosinophil chemotactic cytokine"
/protein_id="AAH11134.1"
/db_xref="GI:15029822"
/translation="MVSTQNRQTFITSVIKPLROYGFDGLDLDWEYFGSRGSPPODK
HLFTLVKEMREAFQEAEIESNRPLMVTAAVAGGISNIOAGYEIPELSKYLDFTHVM
TVDLHSGVEGTGENSPLYKYPTETGSNAYLVNVDYVMYKNGAPAEKLIYGFPEYG
HTFLENSDNGIGAPTSGDGPGAPYTRQAGFWAYEICTFLRSGATVWDASQVVPY
AYKANELGYDNISFSKAKWLQKNFGGMIAIDLDLDDTGSFCDQKPELTSTLN
KALGISTGCTAPDVPSEPTTPPGSGSGSGSGSGSGSGGFCADKADGLYPAVDNRN
AFQCGINTYQHQAGLVDFDTCNCNWP"

CDS

BASE COUNT 393 a 382 c 365 g 398 t
ORIGIN

Alignment Scores:

Pred. No.: 6.9e-92 Length: 1538
Score: 1309.50 Matches: 236
Percent Similarity: 67.33% Conservative: 71
Best Local Similarity: 51.75% Mismatches: 128
Query Match: 54.61% Indels: 21
DB: 10 Gaps: 3

US-10-004-219B-10 (1-445) x BC011134 (1-1538)

QY 3 LeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPheLeuPro 22
Db 68 CTGATATGCTATTTTCCACCACTGGGCCCGCCAGTATCGCCAGGCTCTGGGAGCTTCAAGCT 127

QY 23 LysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMetThrAsn 42
DB 128 GATGACATTAAACCCCTCGCTGTACTACCTACCTATGATCCTTTGCTGGGATGAGA 187
QY 43 HisGlnLeuSerThrGlnTyrPheAsnAspGluThrLeuTyrGlnGluPheAsnGlyLeu 62
DB 188 AATGAGATCACCAACCATAGAAATGGAATGATGATCTCTCTATPAAAGCTTTCATGACTTG 247
QY 63 LysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTyrAsnPheGlyThr 82
DB 248 AAAACAGAACAGCAACTGAAACCCCTCTGGCAATGGAGGCTGGACCTTGGAACT 307
QY 83 GlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsnSerAla 102
DB 308 GCTCCTTTCACTACCATGCTTTCCACTTCTCAGAACCGCCAGACCTTCATTACCTCAGTC 367
QY 103 IleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTyrGluTyrProGly 122
DB 368 ATCAAAATTTCTCGTCAGTATGGGTTGATGACTGGACCTGGACTGGGAATACCCAGGC 427
QY 123 SerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAspLeuAla 142
DB 428 TCAGCTGGAGGCCCTCTCAGACAGCAATCTCTTCACCTGCTCTCTGGTGAAGAAATGGCT 487
QY 143 AsnAlaPheGlnGlnGlnThrSerGlyLysGluArgLeuLeuLeuSerAlaAla 162
DB 488 GAAGCTTTTGAGCAGGAGCTATTGAGACCAACAGGCCAGACCTGATGCTTACTGCTGCT 547
QY 163 ValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAlaGlnAsn 182
DB 548 GTAGCTGGTGGGATTTCCAACTATCCAGGCTGGCTATGAGATCCCTGAACTTCTTAAGTAC 607
QY 183 LeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTyrGlnLysValThrGly 202
DB 608 CTGGATTCATCATGTCATGATCATATGATGATCCATGCTCTCTGGAGGGCTACTGCG 667
QY 203 HisAsnSerProLeuTyrLysArgGlnGluSerGlyValAlaAlaSerLeuAsnVal 222
DB 668 GAGATAGTCTCTTTACAAATACCTACTGAGACTGGTAGCAATGCCTTACCTCAATGTG 727
QY 223 AspAlaAlaValGlnGlnThrLeuGlnLysGlyThrProAlaSerLysLeuIleGly 242
DB 728 GATTATGTCATGAACCTATTGGGAAGAACATGAGGCCCCAGCTGAGAACTCATTTGTGGA 787
QY 243 MetProThrTyrGlyArgSerPheThrLeuAlaSerSerAspThrArgValGlyAla 262
DB 788 TTCCAGAGTAGTGGACACACCTTCATCTCTGAGAAACCCCTCTGATAATGGAATTTGCTGC 847
QY 263 ProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyMetLeuAlaTyr 282
DB 848 CTTACTCTGTGTGATGGCCCTGCTGGCCCTATACCAGACAGGCTGGGTTCTGGGGCTAC 907
QY 283 TyrGluValCysSerTrp-----LysGlyAlaThrLysGlnArgIleGlnAspGlnLys 300
DB 908 TATGAGATTGACCTTTCTGAGAAGTGGAGCCACTGAGGCTCTGGGATGCCCTCCCAAGAA 967
QY 301 ValProTyrIlePheArgAspAsnGlnTyrValGlyPheAspValGluSerPheLys 320
DB 968 GTGCCCTATGCTTAAGCCACAGAGTGGCTTGGCTATGACAAATATCAAGAGCTTCAGT 1027
QY 321 ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyGlyAlaMetValTrpAlaLeuAsp 340
DB 1028 GTTAAGGCTCAGTGGCTTAAGCAGAACAAATTTGGAGGTCATGATGATCTGGGCCATTGAC 1087
QY 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeu 360
DB 1088 CTTGATGACTTCACTGGCTCTTCTGTGATCAGGGAATAATTTCTCTGACTTCTACTTTG 1147
QY 361 ArgGlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLys 380
DB 1148 AACAAAGCCCTGGCATATTCACACTGAAGGTTGACAGCTCCTGACGTG----- 1195
QY 381 ProGlyGlnProSerGluPro-----Glu 388

DB 1196 -----CCTCCGAGCCAGTGACTACTCTCCAGGAAGTGGAGTGGGGTGGAAAGC 1246
QY 389 HisGlyProSerProGlyGlnAspThrPheCysGlnGlyLysAlaAspGlyLeuTyrPro 408
DB 1247 TCCGGAGGAAGCTCTGGAGGAGGAGTGTGTGTCGCAAGCAGATGGGCTCTACCCCT 1306
QY 409 AsnProArgGluArgSerPheTyrSerCysAlaAlaGlyArgLeuPheGlnGlnSer 428
DB 1307 GTGGCAGATGACAGAAATGCTTTTGGCAGTGCATCAATGGAATCACATACCAGGAGCAT 1366
QY 429 CysProThrGlyLeuValPheSerAsnSerCysLysCysCysThrTyr 444
DB 1367 TGTCAAGCAGGCGCTGTTTGTGATACCAGCTGTAATTTGCTGCAACTGG 1414
RESULT 15
BC034548 1557 bp mRNA linear ROD 07-AUG-2002
LOCUS Mus musculus, chitinase, acidic, clone MGC:19045 IMAGE:4189080,
DEFINITION mRNA, complete cds.
ACCESSION BC034548.1 GI:21961190
VERSION BC034548.1
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1557)
Strausberg, R.
Direct Submission
AUTHORS Submitted (24-JUL-2002) National Institutes of Health, Mammalian
TITLE Gene Collection (MGC), Cancer Genomics Office, National Cancer
JOURNAL Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Guanaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 24 Row: h Column: 16
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not
identity to protein.
FEATURES
Location/Qualifiers
1. 1557
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="FVB/N"
/clone="MGC:19045 IMAGE:4189080"
/tissue_type="Salivary gland, 10 week old female mouse"
/clone_lib="NCI CGAP_SG2"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
4. 1425
/codon_start=1
/product="chitinase, acidic"
/protein_id="AAH34548.1"
/db_xref="GI:21961191"
/db_xref="LocusID:81600"
/translation="MAKLLVLTGLALLNAQLGSAYNLCYFTNWAQYRPGLGSPKPD
DINPLCLTHLYAFAGMNNETITWENDVTLYKAFNLDKRNRSKLTLLAIGWNFG

TAPFTTMTSTQNRQTFITTSVIKFLRQYFGDLGLDWEYPSGRSGSPQDKHLFTVLVK
ENREAEQRAIENSRRLMTAVAGGIGNIOAGYEIPBLSKYLDPIHVMYTHLHGSW
EYVTGNSPLKYKPTETGNAVLNVYVMYKNGNAPAEKLIIVGPEYGHFTFILRNP
SDNGICAPTSQDGPAGPYRQAGFWAYEICFLRSGATEVMDASQEVYAVKANEWL
GIDNIKFSVSKQWMLKQNNFNGMIWAILDFTGSCFOGKFPPLTSLUNKALGISTE
GCTAPDVPEPVPTTPGSG
ITVQHCQAGLVFDTSCNCCNP"

BASE COUNT 410 a 382 c 366 g 399 t
ORIGIN

Alignment Scores:

Pred. No.: 7, 01e-92 Length: 1557
Score: 1309.50 Matches: 236
Percent Similarity: 67.32% Conservative: 71
Best Local Similarity: 51.75% Mismatches: 128
Query Match: 54.61% Indels: 21
DB: 10 Gaps: 3

US-10-004-219B-10 (1-445) x BC034548 (1-1557)

QY 3 LeuValCysThrPheThrAsnTPalaGlnTyrArgGlnGlyGluAlaAArgPheLeuPro 22
Db 73 CTGATATGCTATTTCACCAACTGGGCCCGAGTATCGGCCAGGCTCGGGGAGCTTCAAGCCT 132
QY 23 LysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMetThrAsn 42
Db 133 GATGACATTAAACCCCTCGCTGTGTACTACCTGATCTATGCTTTCTGGGATGCAGAAC 192
QY 43 HisGlnLeuSerThrThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsnGlyLeu 62
Db 193 AATGAGATCACCACCATAGAAATGGAATGATGTACTCTCTATAAAGCTTTCAATGACTTG 252
QY 63 LysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTyrAsnPheGlyThr 82
Db 253 AAAACAGGACGACCAACTGAAACCCCTCGCAATTTGGAGCTTGGAACTTTGGAAT 312
QY 83 GlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsnSerAla 102
Db 313 GCTCCTTTCACTACCATGGTTTCCACTTCTCAGAACCGCCAGACCTTCATTACCTCAGTC 372
QY 103 IleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyrProGly 122
Db 373 ATCAAAATTTCTGCGTCAGTATGGGTTTGTAGGACTGGACCTGGACCTGGGAATACCCAGGC 432
QY 123 SerGlnGlySerProAlaValAspLysGluAArgPheThrThrLeuValGlnAspLeuAla 142
Db 433 TCAGTGGGAGCCCTCTCAGGACAGCATCTCTTCTACTGTCTCGTGGAAGGAATGCGT 492
QY 143 AsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuLeuSerAlaAla 162
Db 493 GAAGCTTTTGAGCAGGAGGCTATTGAGAGCAACAGGCCAGACTGATGGTTACTGTGCT 552
QY 163 ValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAlaGlnAsn 182
Db 553 GTAGCTGGTGGGATTTCCAAACATCCAGGCTGGGTATGAGATCCCTGAACTTTCTAAGTAC 612
QY 183 LeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysValThrGly 202
Db 613 CTGATTTTATCCATGTATGATATGACTCTCATGGCTCCCTGGGAGGGCTACACTGGG 672
QY 203 HisAsnSerProLeuTyrLysArgGlnGluSerGlyAlaAlaAlaSerLeuAsnVal 222
Db 673 GAGAAATAGTCTCTTTTACAAATACCTTACTGAGACTGGTAGCAATGCCTACCTCAATGTG 732
QY 223 AspAlaAlaValGlnGlnThrLeuGlnLysGlyThrProAlaSerLysLeuIleLeuGly 242
Db 733 GATTATGTATGTAACATATTGGAAGAACAAATGGAGCCCGCCAGCTGAGAAGCTCATTTGTGA 792
QY 243 MetProThrTyrGlyArgSerPheThrLeuAlaSerSerAspThrArgValGlyAla 262
Db 793 TTCCCAGAGTATGACACACCTTCCTCCTGAGAAACCCCTCTCTATTAATGGAATGGTGCC 852
QY 263 ProAlaThrGlySerGlyThrProGlyThrProPheThrLysGlyGlyMetLeuAlaTyr 282

Search completed: June 29, 2003, 23:53:45

Job time : 2918.5 secs

Db 853 CCTACCTCTGTGTGATGGCCCTGCTGGGCCCTATACCAGACAGGCTGGGTCTTGGGCCTAC 912
QY 283 TyrGluValCysSerTrp-----LysGlyAlaThrLysGlnArgIleGlnAspGlnLys 300
Db 913 TATGAGATTTGACCTTTCTTGAAAGTGGAGCCACTGAGGCTCTGGGATGGCTCCCAAGAA 972
QY 301 ValProTyrIlePheArgAspAsnGlnTrpValGlyPheAspAspValGluSerPheLys 320
Db 973 GTGCCCTATGCCATAAGGCCAACGAGTGGCTTGCTATGACAATATCAAGAGCTTCAGT 1032
QY 321 ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyAlaMetValTrpAlaLeuAsp 340
Db 1033 GTTAAGGCTCAGTGGCTTAAGCAGAACAAATTTTGGAGGTGCCATGATCTGGGCCATTGAC 1092
QY 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeu 360
Db 1093 CTGTAGACTTTCACCTGGCTCTTTCTGTGATCAGGGAATAATTTCTCTGACTTCTACTTTG 1152
QY 361 ArgGlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLys 380
Db 1153 AACAAAGCCCTTGGCATATCCACTGAAGTTGCACAGCTCCTGACGTG----- 1200
QY 381 ProGlyGlnProSerGluPro-----Glu 388
Db 1201 -----CCTTCCGAGCCAGTACTCTCTCCAGGAAGTGGAGTGGGGTGGGAAGC 1251
QY 389 HisGlyProSerProGlyGlnAspThrPheCysGlnGlyLysAlaAspGlyLeuTyrPro 408
Db 1252 TCCGGAGGAGCTCTGGAGGCGAGTGATCTGTGCCGCAAAAGCAGATGGCCTCTACCT 1311
QY 409 AsnProArgGluArgSerSerPheTyrSerCysAlaAlaGlyArgLeuPheGlnGlnSer 428
Db 1312 GTGGCAGATGACAGAAATGCTTTTGGCAGTGCATCAATGGAATCACATACCAGCAGCAT 1371
QY 429 CysProThrGlyLeuValPheSerAsnSerCysLysCysThrTrp 444
Db 1372 TGTCAAGCAGGCGCTTGTTTTGTATACCAGCTGTAAATGTCTGCAACTGG 1419

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 29, 2003, 21:02:46 ; Search time 244.223 Seconds
(without alignments)
4103.372 Million cell updates/sec

Title: US-10-004-219B-10

Perfect score: 2398

Sequence: 1 AKLVCFYTNWAQRQGEARF.....QQSCPTGLVFSNCKCCTWN 445

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xl
-Q/Cgn2_1/USPTO_spool/US10004219/runat_24062003_160227_8151/app_query.fasta_1.1877
-DB=N Geneseq_101002 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptio -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10004219 @CIGN 1.1.626 @runat_24062003_160227_8151 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq_101002:.*
1: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:.*
2: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:.*
3: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:.*
4: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:.*
5: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:.*
6: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:.*
7: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:.*
8: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:.*
9: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:.*
10: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:.*
11: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:.*
12: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:.*
13: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:.*
14: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:.*
15: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:.*
16: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:.*
17: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:.*
18: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:.*
19: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:.*
20: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:.*
21: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:.*
22: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:.*
23: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:.*
24: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2398	100.0	1636	19	AAV10435	Human chitinase cl
2	2398	100.0	1636	20	AAZ21847	MO-218 clone of hu
3	2398	100.0	1636	22	AAD03759	Human chitinase CD
4	2398	100.0	1636	24	ABL57380	Human chitinase CD
5	2394	99.8	1643	18	AAT50833	Human chitinase CD
6	2392	99.7	1656	19	AAV10436	Human chitinase cl
7	2392	99.7	1656	20	AAZ21848	MO-13B clone of hu
8	2392	99.7	1656	22	AAD03760	Human chitinase CD
9	2392	99.7	1656	24	ABL57381	Human chitinase CD
10	2391	99.7	1768	18	AAT89181	Human chitotriosid
11	2372	98.9	1637	18	AAT89180	Human chitotriosid
12	2348.5	97.9	1713	18	AAT50834	Human chitinase CD
13	1323	55.2	1368	22	AAH42013	Disease treatment
14	1323	55.2	1678	22	AAH42025	Disease treatment
15	1072.5	44.7	1369	24	ABN59993	Novel human coding
16	1061	44.2	1594	19	AAV21688	DNA encoding a hum
17	1053	43.9	1432	19	AAV21689	DNA encoding a hum
18	1053	43.9	1594	19	AAV21687	DNA encoding a hum
19	1046.5	43.6	1433	19	AAV13925	Human cartilage gp
20	1046.5	43.6	1496	18	AAT97127	Human cartilage gl
21	1046.5	43.6	1526	19	AAV13926	Human cartilage gp
22	1035.5	43.2	1152	18	AAT99452	Bovine whey protei
23	1030	43.0	1681	16	AAQ85245	YKL-40 gene. Homo
24	1030	43.0	1681	21	AAZ94901	Human cancer marke
25	1015.5	42.3	2346	23	ABL03597	Drosophila melanog
26	996	41.5	1391	24	AA562615	cDNA sequence #402
27	975.5	40.7	17368	23	ABL03596	Drosophila melanog
28	967.5	40.3	1469	22	AAH42023	Disease treatment
29	964.5	40.2	1994	16	AAQ90442	Bovine oviduct spe
30	951	39.7	1474	22	AAH23078	Osteoarthritis tis
31	942	39.3	2504	16	AAQ90443	Murine oviduct spe
32	922	38.4	2366	16	AAQ90444	Hamster oviduct sp
33	919.5	38.3	13497	23	ABL03557	Drosophila melanog
34	848	35.4	1433	23	ABL15099	Drosophila melanog
35	844.5	35.2	16489	23	ABL03556	Drosophila melanog
36	830	34.6	3552	23	ABL15098	Drosophila melanog
37	824	34.4	2452	18	AAT62557	Manduca sexta larv
38	797	33.2	966	9	AAH81756	Gene encoding poly
39	786.5	32.8	2089	23	ABL29841	Drosophila melanog
40	762	31.8	1470	21	AZ38589	D. pteronyssius 98
41	762	31.8	1470	21	AZ38590	D. pteronyssius 98
42	762	31.8	1470	24	ABK69585	Der HMW-map polynu
43	762	31.8	1470	24	ABK69586	Der HMW-map polynu
44	762	31.8	1527	21	AZ38587	D. pteronyssius 98
45	762	31.8	1527	21	AZ38588	D. pteronyssius 98

ALIGNMENTS

RESULT 1
AAV10435
ID AAV10435 standard; cDNA; 1636 BP.
XX
AC AAV10435;
XX
XX
DT 15-JUN-1998 (first entry)
XX
DE Human chitinase clone MO-218 CDNA.

XX Chitinase; human; fungal infection; immunogen; diagnosis; treatment;
XX Gaucher's disease; transgenic; detection; hybridisation; antifungal;
KW rheumatoid arthritis; overexpression; extracellular matrix; ss.
XX Homo sapiens.
OS
XX
FH Key
FT CDS

Location/Qualifiers
2..1402

FT sig_peptide /*tag= a
 FT 2..64
 FT mat_peptide /*tag= b
 FT 65..1399
 FT /*tag= c
 FT /product= chitinase
 FT /note= "from clone MO-218"
 XX
 PN W09747752-A1.

PD 18-DEC-1997.

XX 16-JUN-1997; 97WO-US10460.
 XX 14-JUN-1996; 96US-0663618.
 PA (ICOS-) ICOS CORP.

XX Gray PW;

DR WPI; 1998-052316/05.

DR P-PSDB; AAW40259.

XX Nucleic acids encoding human chitinase - useful as antifungal
 PT agents, especially in combination with other antifungals
 PT
 XX Claim 3; Page 38-40; 63pp; English.

XX This sequence encodes a novel human chitinase isolated from clone MO-218.
 CC Chitinases are useful for treating or preventing fungal infection and
 CC as immunogens for generating antibodies which are used to purify, detect
 CC and quantify chitinases, e.g. for diagnosis of Gaucher's disease. The
 CC nucleic acid sequence of the chitinase is also useful as a probe to
 CC identify and isolate genomic DNA encoding chitinases or similar proteins,
 CC or cells expressing them or to generate transgenic ('knockout') rodents.
 CC It can also be used in hybridisation assays and to detect genetic
 CC alterations in the chitinase gene related to disease. Agents that inhibit
 CC this protein may be useful in treatment of Gaucher's disease and
 CC rheumatoid arthritis, where overexpression of the protein can damage
 CC the extracellular matrix. Chitinase also improves the activity of other
 CC antifungal agents and may allow a reduction in the dose of such agents,
 CC and thus of their side effects.

SQ Sequence 1636 BP; 361 A; 491 C; 440 G; 344 T; 0 other;

Alignment Scores:

Pred. No.: 8-43e-201 Length: 1636
 Score: 2398.00 Matches: 445
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 19 Gaps: 0

US-10-004-219B-10 (1-445) x AAV10435 (1-1636)

QY 1 AlaLysLeuValCysTyrPheThrAsnTTPAlaGlnTyrArgGlnGlyGluAlaArgPhe 20
 DB 65 GCAAACTGGTCTGCTTCTTCCACTGGGCCAGTACAGAGGGGAGGCTCGCTTC 124
 QY 21 LeuProLysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet 40
 DB 125 CTGCCCAAGGACTTGGACCCAGCCTTTCACCCACCTCATCTACGCTTCGCTGGCATG 184
 QY 41 ThrAsnHisGlnLeuSerThrThrGluTTPAsnAspGluThrLeuTyrGlnGluPheAsn 60
 DB 185 ACCAACCCACCGCTGAGCACCCTGAGTGGGAATGACGAGACTCTCTACCGAGAGTTCAAT 244
 QY 61 GlyLeuLysLysMetAsnProLysLeuLysThrLeuAlaIleGlyGlyTTPAsnPhe 80
 DB 245 GGCCTGAAAGAGATGATATCCCAAGCTGAGAGACCTGTTAGCCATCGAGGCTGGAAATTC 304
 QY 81 GlyThrGlnLysPheThrAspMetValAlaThrAlaAsnArgGlnThrPheValAsn 100

DB 305 GGCACCTCAGAGTTCACAGATATGTCAGCCAGCGCAACAACACCTGTCAGACCTTTGTCAAC 364
 QY 101 SerAlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyr 120
 DB 365 TCGGCCATCAGATTCTCGCAATACAGCTTTCAGCCCTTGACCTTGACCTGGGAGTAC 424
 QY 121 ProGlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAsp 140
 DB 425 CCAGGAGCCAGGGAGCCCTGCGGTAGACAGAGGCGCTTCACAAACCTGGTACAGGAC 484
 QY 141 LeuAlaAsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuSer 160
 DB 485 TTGGCCAATGCCCTTCAGCAGGAGCCAGACCTCAGGAGGAAACGCGCTTCTTCTAGT 544
 QY 161 AlaAlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAla 180
 DB 545 GCAGCGGTTCCAGCTGGGCGAGACCTATGTGGATGCTGGATACGAGGTGGACAAAATCGCC 604
 QY 181 GlnAsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysVal 200
 DB 605 CAGAACCTGGATTGTTCACCTTATGGCTTACGACTTCATGGCTCTTGGGAGAGGTC 664
 QY 201 ThrGlyHisAsnSerProLeuTyrLysArgGlnGluGluSerGlyAlaAlaSerLeu 220
 DB 665 ACGGACATACAGCCCTCTTACAAGAGGCAAGAGAGAGTGGTGCAGCAGCCAGCCTC 724
 QY 221 AsnValAspAlaAlaValGlnGlnThrLeuGlnLysGlyThrProAlaSerLysLeuIle 240
 DB 725 AACGTGGATGCTGTGTGCACAGTGGCTGCAAGAGGGAGCCCTGCCAGCAAGCTGATC 784
 QY 241 LeuGlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgVal 260
 DB 785 CTTGGCATGCCCTACCTACGAGCGCTCTTACACCTGGCTCTCATCAGACACCGAGTG 844
 QY 261 GlyAlaProAlaThrGlySerGlyThrProGlyProPheThrLysGlyGlyGlyMetLeu 280
 DB 845 GGGGCCCCAGCCACAGGGTCTGGCACTCCAGGCCCTTCCACCAAGAGAGGAGGATGCTG 904
 QY 281 AlaTyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLys 300
 DB 905 GCCTACTATGAAGTCTGCTCTCGAAGGGGGCCACCAACAGAGAATCCAGGATCAGAAG 964
 QY 301 ValProTyrIlePheArgAspAsnGlnTrpValGlyPheAspValGluSerPheLys 320
 DB 965 GTGCCCTACATCTTCGGGACACACAGTGGTGGCTTTGATGATGGAGAGCTTCANA 1024
 QY 321 ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyGlyAlaMetValTrpAlaLeuAsp 340
 DB 1025 ACCAAGGTCAGCTATCTGAAGCAGAAGGAGCTGGGGGGGCCCATGGTCTGGGCACTGGAC 1084
 QY 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProIleGlnThrLeu 360
 DB 1085 TTAGATGACTTTGGCGGCTTCTCTGCAACAGGCGCGATACCCCTCATCCAGAGCTTA 1144
 QY 361 ArgGlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLys 380
 DB 1145 CGGCAAGAACTGAGTCTTCCATCTTGCCTTCAGGACCCCGAGAGCTTGAAGTTCANAA 1204
 QY 381 ProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGln 400
 DB 1205 CCAGGTCAGCCCTCTGAACCTGAGCATGGCCCGCCCTGGGCAAGACACGCTTCTGCCAG 1264
 QY 401 GlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyrSerCysAla 420
 DB 1265 GGCNAAGCTGATGGGCTCTATCCCAATCTCTGGGAACGGTCCAGCTTCTACAGCTGTGCA 1324
 QY 421 AlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerAsnSerCysLys 440
 DB 1325 GCGGGGCGGCTGTTCACGCAAGCTGCCCGACAGCGCTGGTGTTCAGCAACTCTCTCCANA 1384
 QY 441 CysCysThrTrpAsn 445
 DB 1385 TGCTGCACCTGGAAT 1399

RESULT 2

AAZ21847

ID AAZ21847 standard; DNA; 1636 BP.

XX AC AAZ21847;

XX DT 10-DEC-1999 (first entry)

XX DE MO-218 clone of human Chitinase, with noncoding 5'/3' regions.

XX KW chitin; fungal infection; immunocompromised; AIDS; chemotherapy;

XX KW organ transplant; parasite; chitin-binding; allele; vector;

XX KW truncated protein; ds.

XX OS Homo sapiens.

XX FH Key

XX CDS

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

Location/Qualifiers

2..1402

/*tag= a

/product= Human_Chitinase

2..65

/*tag= b

/product= Signal_peptide

66..1402

/*tag= c

/product= Mature_protein

W09946390-A1.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Novel chitin-binding fragments of human chitinase used to treat fungal infections in animals

Example 1; Page 55-57; 83pp; English.

This is the nucleotide sequence of an allelic form of the human chitinase enzyme, which is capable of degrading Chitin (a linear homo polymer of beta-1,4-linked N-acetylglucosamine residues). Chitinase fragments can be used to screen for proteins or other molecules that specifically bind to the chitin-binding domain of human chitinase or that modulate its activity. These compounds are useful for immunization, as well as for purifying chitinase, as well as for detection and quantification of chitinase. Polynucleotide fragments of the invention are useful as a source of probes and primers, and to express the proteins recombinantly. The chitinase fragments, when conjugated to antifungal compounds, are used to treat animals, especially humans, infected with chitin-containing parasites such as fungi. Fungal infection treated include candidiasis, aspergillosis, mucormycosis, blastomycosis, paracoccidioidomycosis, coccidioidomycosis, histoplasmosis, cryptococcosis, chromoblastomycosis, sporotrichosis, and dermatophytoses. Chitin can be degraded by the enzyme chitinase. Use of whole chitinase protein for treating infections, especially fungal infections, is problematic. In view of the increasing incidents of life-threatening fungal infection in e.g. immunocompromised individuals, there exists a need for identifying new compounds for treating fungal infection. The chitin-binding fragments of the present invention provide this need.

Sequence 1636 BP; 361 A; 491 C; 440 G; 344 T; 0 other;

Alignment Scores:

Pred. No.: 8.43e-201 Length: 1636
Score: 2398.00 Matches: 445
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 20 Gaps: 0

US-10-004-219B-10 (1-445) x AAZ21847 (1-1636)

QY 1 AlaLysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPhe 20
DB 65 GCAAACTGGTCTGCTACTTACCACCTGGGCCAGTACAGACAGGGGAGGCTCGCTTC 124
QY 21 LeuProLysAspLeuAaspProSerLeuCysThrHisLeuLeuLeuAlaGlyMet 40
DB 125 CTGCCCAAGGACTTGGACCCCGCCCTTTGCACCCACCTCATCTAGCCCTTCGCTGGCATG 184
QY 41 ThrAsnHisGlnLeuSerThrThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsn 60
DB 185 ACCAACCCACAGCTGAGCACCACTGAGTGAATGACGAGACTCTCTACCCAGGAGTTCAAT 244
QY 61 GlyLeuLysLysMetAsnProLysLeuLysThrLeuLeuAlaLeuGlyTrpAsnPhe 80
DB 245 GGCCTGAAGAAGATGANTCCCAAGCTGAAGACCCCTGTTAGCCATCGAGGCTGGAAATTC 304
QY 81 GlyThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsn 100
DB 305 GGCACCTCAGAGTTCCAGATATGTTAGCCAGCCGACACCCGTCAGACCTTTGTCAAC 364
QY 101 SerAlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyr 120
DB 365 TCGGCCATCAGGTTTCTGGCAATACAGCTTTGACGGCCTTGACTTGGAGTAC 424
QY 121 ProGlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAsp 140
DB 425 CCAGGAAGCCAGGGAGGCCCTGCGGTAGACAGGAGCGCTTCACAAACCTTGTACAGGAC 484
QY 141 LeuAlaAsnAlaPheGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 160
DB 485 TTGGCAATGCTTCCAGGAGAACCCAGACCTCAGGAGAGAACCGCTTCTTCTGAGT 544
QY 161 AlaAlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAla 180
DB 545 GCAGCGGTTCCAGCTGGGAGGAGGCTATGTGGATGCTGGATACGAGGTGACAAATCGCC 604
QY 181 GlnAsnLeuAaspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysVal 200
DB 605 CAGAACTGGGATTTTGTCAACCTTATGGCTACGACTTCCATGGCTCTTGGGAGAGGTC 664
QY 201 ThrGlyHisAsnSerProLeuTyrLysArgGlnGluSerGlyValAlaAlaSerLeu 220
DB 665 ACGGACATTAACAGCCCTCTTACAAAGAGGCAAGAGAGTGGTGCAGCAGCCAGCTTC 724
QY 221 AsnValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuIle 240
DB 725 AACGTGGATGCTGCTGCAACAGTGGCTGCAGAGGGGAGCCCTGCCAGCAAGCTGATC 784
QY 241 LeuGlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerThrArgVal 260
DB 785 CTTGGCATGCTTACCTACGAGGAGCTCTTCACTAGCCCTCTCATCAGACACCCAGAGTG 844
QY 261 GlyAlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyGlyMetLeu 280
DB 845 GGGGGCCCGCCAGGAGGCTGCGCACTCCAGGCCCTTCCACCAAGAGAGGAGGAGTCTG 904
QY 281 AlaTyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLys 300
DB 905 GCCTACTATGAGTCTGCTCTCTGGAGGGGGCCCAACACAGAGATCCAGGATCAGAG 964
QY 301 ValProTyrIlePheArgAspAsnGlnTrpValGlyPheAspValGluSerPheLys 320
DB 965 GTGCCCTACATCTTCGGGACCAACGAGTGGTGGGCTTTGATGATGTGGAGAGCTTCAAA 1024

QY 321 ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyGlyAlaMetValTTPAlaLeuAsp 340
 DB 1025 ACCAGGTGAGTATCTGAGAGGAGAGGAGTCTGGGGGGCCATGCTCTGGGCACTGGAC 1084

QY 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuLeuGlnThrLeu 360
 DB 1085 TTAGATGACTTTGCGGCTTCTCTGCAACAGGCGCGATACCCCTCATCCAGACGCTA 1144

QY 361 ArgGlnLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLys 380
 DB 1145 CGCAGGAACTGAGTCTTCATCTTCCCTTCAGGCAACCCAGAGTTGAAGTTCCAAA 1204

QY 381 ProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGln 400
 DB 1205 CAGGTGAGCCCTCTGACCTGAGCATGGCCCGCCCTGGACAGACAGCTTCTGCCAG 1264

QY 401 GlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyrSerCysAla 420
 DB 1265 GGCAGAGCTGATGGCTCTATCCCAATCTCTCGGACGCTCCAGCTTCTACAGCTGTGCA 1324

QY 421 AlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerAsnSerCysLys 440
 DB 1325 GCGGGCGGCTGTCTCCAGCAAGCTGCCGACAGGCTGTGTTCAGCAACTCTCCAAA 1384

QY 441 CysCysThrTTPAsn 445
 DB 1385 TGCTGCACTGGAT 1399

RESULT 3
 AAD03759
 ID AAD03759 standard; cDNA; 1636 BP.
 AC AAD03759;
 XX
 DT 19-JUN-2001 (first entry)
 DE Human chitinase cDNA from clone pMO-218.
 KW Human; antifungal; chitinase; immunoglobulin; Ig; therapy;
 KW fungal infection; candidiasis; aspergillosis; coccidioidomycosis;
 KW blastomycosis; paracoccidioidomycosis; histoplasmosis; cryptococcosis;
 KW chromoblastomycosis; sporotrichosis; mucormycosis; dermatophytoses;
 KW clone pMO-218; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 2..1402
 FT /*tag= a
 FT /product= "Human chitinase from clone pMO-218"
 FT sig_peptide 2..64
 FT /*tag= b
 FT mat_peptide 65..1399
 FT /*tag= c
 FT /product= "Human mature chitinase from clone pMO-218"
 XX
 PN WO200123430-A2.
 XX
 PD 05-APR-2001.
 XX
 PF 28-SEP-2000; 2000WO-US26960.
 XX
 PR 30-SEP-1999; 99US-0409918.
 XX
 PA (ICOS-) ICOS CORP.
 XX
 PI Allison DS, Dietsch GN, Gray PW, Shaw KD, Steiner BH;
 XX
 DR WPI: 2001-266141/27.
 DR P-PSDB; AAE00432.
 XX
 FT Novel chitinase immunoglobulin fusion product, useful for treating

PT fungal infections and reducing the amount of a non-chitinase antifungal
 agent needed for the treatment -
 XX
 PS Claim 2; Page 29-31; 39pp; English.
 XX
 CC The present invention relates to a chitinase immunoglobulin (Ig) fusion
 product, comprising a human chitinase fused to at least a portion of an
 immunoglobulin chain. The fusion product is useful for treating fungal
 infections (mycoses) such as candidiasis, aspergillosis, blastomycosis,
 coccidioidomycosis, paracoccidioidomycosis, histoplasmosis, mucormycosis,
 cryptococcosis, chromoblastomycosis, sporotrichosis and dermatophytoses.
 The fusion protein is useful for reducing the amount of non-chitinase
 antifungal agent needed to exert an antifungal activity. The fusion
 protein is also useful for preparing a medicament for the prophylactic
 or therapeutic treatment of fungal infections. Chitinase immunoglobulin
 fusion product has unexpectedly improved serum half-life and formulation
 properties. The present sequence is human chitinase cDNA from clone
 pMO-218. Chitinase enzyme degrades chitin which is a homopolymer of
 beta-(1,4)-linked N-acetylglucosamine residues.
 XX
 SQ Sequence 1636 BP; 361 A; 491 C; 440 G; 344 T; 0 other;

Alignment Scores:
 Pred. No.: 8,43e-201 Length: 1636
 Score: 2398.00 Matches: 445
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 22 Gaps: 0

US-10-004-219B-10 (1-445) x AAD03759 (1-1636)

QY 1 AlaLysLeuValCysTyrPheThrAsnTTPAlaGlnTyrArgGlnGlyGluAlaArgPhe 20
 DB 65 GCAAACTGGTCTGCTACTTCCAACTGGGCCCACTAGACAGAGGGAGGCTCGCTTC 124

QY 21 LeuProLysAspLeuAspProSerLeuCysThrHisLeuLeuTyrAlaPheAlaGlyMet 40
 DB 125 CTGCCCAAGGACTTGGACCCCGCCAGCCTTTGCCACCCACCTCATCTACGCTTCGCTGGCATG 184

QY 41 ThrAsnHisGlnLeuSerThrThrGluTTPAsnAspGluThrLeuTyrGlnGluPheAsn 60
 DB 185 ACCAAACCAACAGCTGAGCACCACCTGAGTGGATGACGAGACTCTCTACAGAGGTTCAT 244

QY 61 GlyLeuLysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTTPAsnPhe 80
 DB 245 GGCCTGAGAGAGATGATCCCAAGCTGAGACCCCTGTAGCCATCGAGCTGGAAATTC 304

QY 81 GlyThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsn 100
 DB 305 GGCACCTCAGAGTTCCACAGATATGTAGCCACGGCCCAACACCGTCAGACCTTTGTCAAC 364

QY 101 SerAlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTTPGluTyr 120
 DB 365 TCGGCCATCAGGTTTCTGCGCAATACAGCTTTGACGCTTACGCTTACCTGGAGTAC 424

QY 121 ProGlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAsp 140
 DB 425 CCAGGAAGCCAGGGAGGCTTCCTCCAGCAGGAGCCAGACCTCAGGGAAGGAACGCTTCTCTGAGT 484

QY 141 LeuAlaAsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuSer 160
 DB 485 TTGGCCAATGCTTCCAGCAGGAGCCAGACCTCAGGGAAGGAACGCTTCTCTGAGT 544

QY 161 AlaAlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAla 180
 DB 545 GCAGCGGTTCAGCTGGGCGAGACCTATGTGGATGTGGATACGAGGTGGACAAAATCGCC 604

QY 181 GlnAsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTTPGluLysVal 200
 DB 605 CAGAACCTGGATTTGTCAACCTTATGGCTTACGACTTCCATGGCTCTTGGAGAGGTC 664

QY 201 ThrGlyHisAsnSerProLeuTyrLysArgGlnGluSerGlyAlaAlaSerLeu 220

Db 665 ACAGGACATACAGACCCCTCTACAGAGGCAAGAGAGTGGTGCAGCAGCCAGCCTC 724
Qy 221 AenValAspAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerIleuLeu 240
Db 725 AACGTGGATGCTGCTGTGCAACAGTGGCTGCAAGAGGAGCCCTGCAGCAAGCTGATC 784
Qy 241 LeuGlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerAspThrArgVal 260
Db 785 CTTGGCATGCTTACCTACGAGCGCTCTTACACATGGCTCTCATCAGACACAGAGTG 844
Qy 261 GlyAlaProAlaThrGlySerGlyThrProGlyProPheThrLysGlnGlyMetLeu 280
Db 845 GGGGCCCCCAGCCACAGGCTGCTGCACTCCAGGCCCTTCCACCAAGAGAGGAGTGCTG 904
Qy 281 AlaTyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLys 300
Db 905 GCCTACTATGAAGTCTGCTCTCGGAAGGGGGCCACCAACAGAGAAATCAGGATCAGAG 964
Qy 301 ValProTyrIlePheArgAspAsnGlnTrpValGlyPheAspValGluSerPheLys 320
Db 965 GTGCCCTACATCTTCGGGACACACAGTGGGTGGCTTTGATGATGTGGAGAGCTTCAA 1024
Qy 321 ThrLysValSerTyrIleuLysGlnLysGlyLeuGlyAlaMetValTrpAlaLeuAsp 340
Db 1025 ACCAAGGTGAGTATCTGAAGCAGAGGAGTGGCGGGGCCATGCTCTGGGCACCTGGAC 1084
Qy 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeu 360
Db 1085 TTAGATGACTTTGGCGGCTTCTCTGCAACACAGGGCCGATACCCCTCATCCAGAGCTTA 1144
Qy 361 ArgGlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLys 380
Db 1145 CGGCAGGAAGTCTTCATCTTCCATCTTCCAGGACCCAGAGCTTGAAGTTCCAAA 1204
Qy 381 ProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGln 400
Db 1205 CCAGGTGACGCCCTCTGAACCTGAGCATGCCGCCAGCCCTGGCAAGACACGCTTCTGCCAG 1264
Qy 401 GlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerPheTyrSerCysAla 420
Db 1265 GCAAAAGCTGTATGGGCTCTATCCCAATCTCGGGAACGGTCCAGCTTCTACAGCTGTGCA 1324
Qy 421 AlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerAsnSerCysLys 440
Db 1325 GGGGGGGGCTTTCCAGCAAGAGCTGCCGACAGGCTGGTGTTCAGCAACTCTCTGCCAA 1384
Qy 441 CysCysThrTrpAsn 445
Db 1385 TGCTGCACCTGGAAT 1399

RESULT 4

ABL57380
ID ABL57380 standard; cDNA; 1636 BP.

AC ABL57380;

DT 12-AUG-2002 (first entry)

XX Human chitinase cDNA clone MO-218.

DE Chitinase; enzyme; human; fungicide; antifungal; infection;
KW candidiasis; aspergillosis; coccidioidomycosis; blastomycosis;
KW paracoccidioidomycosis; histoplasmosis; cryptococcosis;
KW Pneumocystis; sporotrichosis; mucormycosis; dermatophytosis;
KW gene; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

FT 2..1402

FT /tag= a

FT /product= "Chitinase"

sig_peptide 2..64
mat_peptide /*tag= b
65..1399
/*tag= c
US6372212-B1.
16-APR-2002.
16-JUN-1997; 97US-0877599.
14-JUN-1996; 96US-0663618.
(ICOS-) ICOS CORP.
Gray PW;
WPI; 2002-442449/47.
P-PSDB; ABB76291.

Co-administering chitinase to improve the effectiveness of fungicidal drugs e.g. amphotericin B or itraconazole, useful for treating fungal infections e.g. candidiasis, coccidioidomycosis and blastomycosis -
Example 1; Column 23-28; 26pp; English.

The present sequence is that of cDNA clone MO-218 (ATCC 98077) encoding human chitinase (see ABB76291). The clone was isolated from a cDNA library prepared from peripheral blood monocyte-derived macrophages following sequence analysis. Also isolated was clone MO-13B (see ABL57381), which contains a single nucleotide difference in the coding region, changing the encoded amino acid at position 81 of the mature protein from glycine to serine. Northern blots showed highest chitinase gene expression in lung and ovary tissues. Expression in lung is consistent with a protective role against pathogenic organisms that contain chitin. The invention provides human chitinase polynucleotides and polypeptides, and materials and methods for the recombinant production of human chitinase products, which are expected to be useful as products for treating fungal infections or for the development of such products. Human chitinase has a synergistic effect on the actions of other fungicides. It can be administered to improve the antifungal activity of a non-chitinase antifungal agent, especially amphotericin B or itraconazole, in the treatment of a fungal infection such as candidiasis, aspergillosis, coccidioidomycosis, blastomycosis, paracoccidioidomycosis, histoplasmosis, cryptococcosis, chromoblastomycosis, sporotrichosis, mucormycosis, dermatophytoses and Pneumocystis infections (all claimed). In particular, the fungal infection involves Candida, Aspergillus and/or Cryptococcus spp., whose growth is not effectively inhibited by contact with human chitinase alone.

SQ Sequence 1636 BP; 361 A; 491 C; 440 G; 344 T; 0 other;

Alignment Scores:

Pred. No.: 8, 43e-201 Length: 1636
Score: 2398.00 Matches: 445
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 24 Gaps: 0

US-10-004-219B-10 (1-445) x ABL57380 (1-1636)

Qy 1 AlaLysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyAlaArgPhe 20
Db 65 GCAAAACTGGTGTCTACTTCCAACTGGGCCCATAGTACAGACAGGGGAGCTCGCTTC 124
Qy 21 LeuProLysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet 40
Db 125 CTGCCCAAGGAGCTTGGACCCCGAGCCCTTTGGACCCCACTCATCTACGCTTCGCGCATG 184
Qy 41 ThrAsnHisGlnLeuSerThrThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsn 60

```
Db 185 ACCAACCCAGCTGAGCACCCTGAGTGAATGACGAGACTCTTACCGAGGATTCAT 244
Qy 61 GVLLeuLysMetAsnProLysLeuLysLeuLeuAlaLeuGlyGlyTyrPheAsn 80
Db 245 GGCCTGAAGAAGATGAATCCAGCTGAAGACCCCTGTATGCCATCCGAGGCTGGAATTC 304
Qy 81 GLYThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsn 100
Db 305 GGCACCTCAGAGTTTCAGATATGTGTAGCCAGCCGCAACCAACCGTCAGACCTTTGTCAAC 364
Qy 101 SerAlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTyrGluTyr 120
Db 365 TCGGCCATCAGGTTCTGGCCAAATACAGCTTTCAGCGGCTTGACCTGACCTGGAGTAC 424
Qy 121 ProGlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAsp 140
Db 425 CCAGGAAGCAGGGAGCCCTGCGGTAGACAGAGCGCTTCAACCCCTGGTACAGGAC 484
Qy 141 LeuAlaAsnAlaPheGlnGlnGlnAlaGlnThrSerGlyLysGluArgLeuLeuLeuSer 160
Db 485 TTGGCCAAATGCCCTTCCAGCAGGAAGCCAGACCTCAGGGAAGAACGCTTCTTCTGAGT 544
Qy 161 AlaAlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAla 180
Db 545 GCAGCGGTTCCAGCTGGCAGACCTATGTGATGCTGATACGAGGTGACAAAATCGCC 604
Qy 181 GlnAsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTyrGluLysVal 200
Db 605 CAGAACTGTGATTTGTCAACCTTATGGCCCTAGCACTTCCATGGCTCTTGGGAAGAGGC 664
Qy 201 ThrGlyHisAsnSerProLeuTyrLysArgGlnGluSerGlyAlaAlaLaserLeu 220
Db 665 ACGGACATTAACAGCCCTCTACAGAGCGCAAGAGAGTGTGTCAGCAGCCAGCCCTC 724
Qy 221 AsnValAspAlaValGlnGlnThrLeuGlnLysGlyThrProAlaSerLysLeuIle 240
Db 725 AACGTGGATGCTGCTGTCAACAGTGGCTGTGACAGAGGGACCCCTGCCAGCAAGCTGATC 784
Qy 241 LeuGlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgVal 260
Db 785 CTGGGATGCTTACCTACGAGCCCTCTTCACTGGCTTCTCATCAGACACAGAGTG 844
Qy 261 GlyAlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyMetLeu 280
Db 845 GGGGCCCCAGCCACAGGCTCTGCACTTCCAGGCCCTTCAACGAAGAGGAGGATGCTG 904
Qy 281 AlaTyrTyrGluValCysSerTyrLysGlyAlaThrLysGlnArgIleGlnAspGlnLys 300
Db 905 GCCTACTATGAGTCTGCTCTGGAAGGGGGCCCAACCAACAGAGAATCCAGGATCAGAAG 964
Qy 301 ValProTyrIlePheArgAspAsnGlnTyrValGlyPheAspValGluSerPheLys 320
Db 965 GTGCCCTACATCTCCGGGCAACCAAGTGGTGGCTTTGATGATGGAGAGCTTCAAA 1024
Qy 321 ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyAlaMetValTyrPalaLeuAsp 340
Db 1025 ACCAAGGTGAGCTATCTGAAGCAGAGGAGCTGGCGGGGCCATGCTTGGGCACCTGGAC 1084
Qy 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeu 360
Db 1085 TTAGATGACTTTTCCGGGCTTCTCTGCAACAGGCGGCGATACCCCTCATCCAGAGCTA 1144
Qy 361 ArgGlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLys 380
Db 1145 CGGCAGAACTGAGTCTTCATACTTTCCTTTCAGGACCCCGAGCTTGAAGTTCCAAA 1204
Qy 381 ProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGln 400
Db 1205 CCAGGTGAGCCCTCTGAACCTGAGCATGGCCCGCCAGCCCTGGACAGACAGCTTCTGCCAG 1264
Qy 401 GlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyrSerCysAla 420
```

```
Db 1265 GCGAAAGCTGATGGGCTCTATCCCAATCTCCGGAACGGTCCAGCTTCTACAGCTGTGCA 1324
Qy 421 AlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerAsnSerCysLys 440
Db 1325 GGGGGGGGCTGTTCAGCAAGGCTGCCGACAGAGCTGTGTTCAGCAACTCTCTGCANA 1384
Qy 441 CysCysThrTyrAsn 445
Db 1385 TGCTGCACCTGGAAT 1399

RESULT 5
AAT50833
ID AAT50833 standard; cDNA; 1643 BP.
XX
AC AAT50833;
XX
DT 24-MAR-1997 (first entry)
XX
DE Human chitinase cDNA clone chi.50.
XX
KW Chitinase; chitotriosidase; chitin; infectious disease;
KW gene therapy; vaccine; lysosomal lipodosis; Gaucher's disease;
KW leishmaniasis; sarcoidosis; X-linked adrenoleukodystrophy;
KW multiple sclerosis; drug delivery; cosmetics; food; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 13..1413
FT sig_peptide /*tag= a
FT mat_peptide /*tag= b
FT /*tag= c
XX
PN WO9640940-A2.
XX
PD 19-DEC-1996.
XX
PF 06-JUN-1996; 96WO-NL00225.
XX
PR 07-JUN-1995; 95US-0486839.
XX
(PNAM ) UNIV AMSTERDAM.
XX
PI Aerts JWFG;
XX
XX WPI; 1997-118698/11.
DR P-PSDB; AAW08584.
XX
PT New human chitinase - used to treat or prevent infection by
PT chitin-contg. pathogens, in diagnosis and as additives to cosmetics,
PT foods, implants etc.
XX
PS Claim 2; Page 39-40; 58pp; English.
XX
CC A cDNA clone (AAT50833), designated chi.50, codes for a 50 kDa human
CC chitinase (AAW08584) that is stable to many proteases, active at pH 3-
CC 8 and up to 50 deg, and stable in the circulation. Clones chi.50
CC and chi.39 (see also AAT50834) were isolated from a human macrophage
CC cDNA library using as probe a partial clone obtd. using primers
CC (see also AAT50835-36) based on a chitotriosidase purified from a
CC type 1 Gaucher disease patient. The 2 cDNA clones are the result
CC of alternative splicing of RNA. Chitinase nucleic acid can be used
CC for large-scale prodn. of recombinant human chitinases, or can be
CC incorporated into a gene therapy vector to treat or prevent
CC infection by chitin-contg. pathogens.
XX
SQ Sequence 1643 BP; 364 A; 490 C; 442 G; 347 T; 0 other;

Alignment Scores: 1.9e-200 Length: 1543
Pred. No.: 2394.00 Matches: 444
Score:
```

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.83% Indels: 0
DB: 18 Gaps: 0

US-10-004-219B-10 (1-445) x AAT50833 (1-1643)

2 LysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPheLeu 21
79 AAACCTGGTCTGCTACTTCAACCACTGGGCCAGTACAGACAGGGGGAGGCTGCTTCCTG 138
22 ProLysAspLeuAspProSerLeuCysThrHisLeulleTyrAlaPheAlaGlyMetThr 41
139 CCCAAGGACTTGGACCCAGCCCTTTGGACCCACCTCATCTAGCCCTTGGCTGGCATGCC 198
42 AsnHisGlnLeuSerThrTrpGlnTrpAsnAspGluThrLeuTyrGlnGluPheAsnGly 61
199 AACCAACAGCTGAGCACCACCTGAGTGGATGACGAGACTCTCTACAGGAGTTCAATGCG 258
62 LeuLysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPheGly 81
259 CTGAAGAGATGAATCCCAAGCTGAAGACCTTGTAGCCATCGGAGGCTGGAAATTCGCG 318
82 ThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsnSer 101
319 ACTCAAGAGTTCAACAGATATGTTAGCCAGCGGCCAACCCGTGACACCTTTGTCAACTCG 378
102 AlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTrpPro 121
379 GCCATCAGGTTTCTCGGCAATACAGCTTTGACGGCTTGTACCTTGGAGTACCCA 438
122 GlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAspLeu 141
439 CGAAGCCAGGGGAGCCCTGCGCTAGCAAGAGAGCGCTTCAACACCTGGTACAGACTTG 498
142 AlaAsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuSerAla 161
499 GCCAATGCTTCCAGCAGGAAGCCAGACCTCAGGGAAGAAACGCTTCTTCTGAGTGCA 558
162 AlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAlaGln 181
559 GCGGTTCAGCTGGGAGACCTATGTGGATGCTGGATAGAGTGGACAAAATCGCCAG 618
182 AsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysValThr 201
619 AACCTGGATTTGTCAACCTTATGGCTTACGACTTCCATGGCTCTTGGGAGAGGTCAAG 678
202 GlyHisAsnSerProLeuTyrLysArgGlnGlnGluSerGlyAlaAlaAspSerLeuAsn 221
679 GGACATAACAGCCCTCTTACAAGAGGCAAGAGAGAGTGTGCGACAGCCAGCCTCAAC 738
222 ValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuIleLeu 241
739 GTGGATGCTGCTGTCAACAGTGGCTGCAAGAGGGGACCCCTGCCAGCAAGCTGATCCTT 798
242 GlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgValGly 261
799 GGCATGCTTACCTACGAGCGCTCTTCACTGGCTCTCTCATCAGACACAGAGTGGGG 858
262 AlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyGlyMetLeuAla 281
859 GCCCCAGGCCACAGGGTCTGGCATCTCCAGGCCCTTCAACAGGAAGAGGGATGCTGGCC 918
282 TyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLysVal 301
919 TACTATGAAGTCTGCTCTCGGAAGGGGGCCCAACAGAGAATCCAGGATCAGAGGTG 978
302 ProTyrIlePheArgAspAsnGlnTrpValGlyPheAspAspValGluSerPheLysThr 321
979 CCTACATCTTCGGGACACCAACAGTGGGTGGCTTTGATGATGGAGAGCTTCAAACC 1038
322 LysValSerTyrLeuLysGlnLysGlyLeuGlyAlaMetValTrpAlaLeuAspLeu 341

Db 1039 AAGGTCACTATCTCAAGACAGAGGAGTGGCGGGGCCATGGTCTGGGCACCTGGACTTA 1098
Qy 342 AspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeuArg 361
Db 1099 GATGACTTTGGCGGCTTCTCTCTCAACCGAGGGCCGATACCCCTCATCCAGCGTACGG 1158
Qy 362 GlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLysPro 381
Db 1159 CAGGAACCTGAGTCTTCCATACCTTGCCTTCAGGCACCCCGAGAGCTTGAAGTTCCAAACCA 1218
Qy 382 GlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGlnGly 401
Db 1219 GGTCAAGCCTCTGAACCTGAGCATGGCCCGCCGCTGGACAGACACCTTCTGCCAGGCG 1278
Qy 402 LysAlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyrSerCysAlaAla 421
Db 1279 AAAGCTGATGGGCTCTATCCCAATCCTCGGACCGGTCCAGCTTCTACAGCTGTGCAGG 1338
Qy 422 GlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerAsnSerCysLysCys 441
Db 1339 GGGCGGCTGTTCCAGCAAGAGCTGCCGACAGGCTGGTGTTCAGCAACTCCTGCAATGC 1398
Qy 442 CysThrTrpAsn 445
Db 1399 TGCACCTGGAAT 1410
RESULT 6
AAV10436
ID AAV10436 standard; cDNA; 1656 BP.
XX AAV10436;
AC AAV10436;
XX XX
DT 15-JUN-1998 (first entry)
XX XX
DE Human chitinase clone MO-13B cDNA.
XX XX
KW Chitinase; human; fungal infection; immunogen; diagnosis; treatment;
KW Gaucher's disease; transgenic; detection; hybridisation; antifungal;
KW rheumatoid arthritis; overexpression; extracellular matrix; ss.
XX XX
OS Homo sapiens.
XX XX
FH Key Location/Qualifiers
FT CDS 27..1427
FT /*tag= a
FT sig_peptide 27..89
FT /*tag= b
FT mat_peptide 90..1424
FT /*tag= c
FT /product= chitinase
FT /note= "from clone MO-13B"
XX XX
PN WO9747752-A1.
XX XX
PD 18-DEC-1997.
XX XX
PF 16-JUN-1997; 97WO-US10460.
XX XX
PR 14-JUN-1996; 96US-0663618.
XX XX
PA (ICOS-) ICOS CORP.
XX XX
PI Gray PW;
XX XX
DR WPI; 1998-052316/05.
DR P-PSDB; AAW40260.
XX XX
PT Nucleic acids encoding human chitinase - useful as antifungal
XX XX
PS agents, especially in combination with other antifungals
XX XX
CC Claim 9; Page 42-44; 63pp; English.
XX XX
CC This sequence encodes a novel human chitinase isolated from clone MO-13B.

CC Chitinases are useful for treating or preventing fungal infection and
 CC as immunogens for generating antibodies which are used to purify, detect
 CC and quantify chitinases, e.g. for diagnosis of Gaucher's disease. The
 CC nucleic acid sequence of the chitinase is also useful as a probe to
 CC identify and isolate genomic DNA encoding chitinases or similar proteins,
 CC or cells expressing them or to generate transgenic ('knockout') rodents.
 CC It can also be used in hybridisation assays and to detect genetic
 CC alterations in the chitinase gene related to disease. Agents that inhibit
 CC this protein may be useful in treatment of Gaucher's disease and
 CC rheumatoid arthritis, where overexpression of the protein can damage
 CC the extracellular matrix. Chitinase also improves the activity of other
 CC antifungal agents and may allow a reduction in the dose of such agents,
 CC and thus of their side effects.

XX
 SQ Sequence 1656 BP; 365 A; 497 C; 447 G; 347 T; 0 other;

Alignment Scores:

Pred. No.: 2,886-200 Length: 1656
 Score: 2392.00 Matches: 444
 Percent Similarity: 99.78% Conservative: 0
 Best Local Similarity: 99.78% Mismatches: 1
 Query Match: 99.75% Indels: 0
 DB: 19 Gaps: 0

US-10-004-219B-10 (1-445) x ARAV10436 (1-1656)

QY 1 AlaValLeuValCysThrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPhe 20
 DB 90 GCAAAACTGCTGCTGCTACCTTCCCACTGGCGCCAGTACAGAGGGGGAGGCTCGCTTC 149
 QY 21 LeuProLysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet 40
 DB 150 CTGCCCAAGGACTTGGACCCCGCCCTTTGCACCCACTCATCGCTTCGTGGCATG 209
 QY 41 ThrAsnHisGlnLeuSerThrThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsn 60
 DB 210 ACCAACCCAGCTGAGCACCCTAGTGGTGAATGACGAGACTCTTACCAGGAGTTCAAT 269
 QY 61 GlyLeuLysMetAsnProLysLeuLysThrLeuAlaIleGlyGlyTrpAsnPhe 80
 DB 270 GGCTTGAAGAAGATGAATCCCAAGCTGAGAGCCCTGTAGCCATCGAGCTGGAAATTC 329
 QY 81 GlyThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsn 100
 DB 330 AGCACTCAGAAGTTACAGATATGTTAGCCAGCCGCAACACCGTCAGACCTTTGTCAAC 389
 QY 101 SerAlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyr 120
 DB 390 TCGGCCATCAGGTTTCTGCGCAATACAGCTTTGACGGCCCTTGACCTTGACCTGGGAGTAC 449
 QY 121 ProGlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAsp 140
 DB 450 CCAGGAAGCCAGGGAGCCCTGCGGTAGAGAGAGCGCTTCAACACCTGGTACAGGAC 509
 QY 141 LeuAlaAsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuSer 160
 DB 510 TTGGCCAATGCTTCCAGCAGGAAGCCAGACCTCAGGAGGAAGAACGCTTCTCTGAGT 569
 QY 161 AlaAlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAla 180
 DB 570 GCAGCGGTTCCAGCTGGGAGACCTATGTGGATGCTGGATACGAGGTGGACAAAATCGCC 629
 QY 181 GlnAsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysVal 200
 DB 630 CAGAACCTGATTTGTGCAACCTTATGGCTACGACTTCATGGCTCTTGGGAGAGGTC 689
 QY 201 ThrGlyHisAsnSerProLeuTyrLysArgGlnGluSerGlyAlaAlaLaserLeu 220
 DB 690 ACGGGACATAACAGCCCTCTTACAGAGGCAAGAGAGAGTGTGTGACGAGCCCTC 749
 QY 221 AsnValAspAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuIle 240
 DB 750 AACCTGGATGCTGCTGTGCAACAGTGGCTGCAAGAGGGGACCCCTGCCAGAGCTGATC 809

QY 241 LeuGlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgVal 260
 DB 810 CTTGGCATGCTTACCTACGAGCGCTCTCTTCCACCTGGCTCTCTCATCAGACACAGAGTG 869
 QY 261 GlyAlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyGlyMetLeu 280
 DB 870 GGGGCCCCCAGCCACAGGCTTGGCACTCAGGCCCTTCCACCAAGGAAGGAGGATGCTG 929
 QY 281 AlaTyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLys 300
 DB 930 GCCTACTATGAAGTCTGCTCTGGAAGGGGGGCCCAACACAGAGAAATCCAGATCAGAAG 989
 QY 301 ValProTyrIlePheArgAspAsnGlnTrpValGlyPheAspAspValGluSerPheLys 320
 DB 990 GTGCCCTACATCTTCCGGAGCAACAGCTGGGTGGCTTGTATGATGTGGAGAGCTTCANA 1049
 QY 321 ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyGlyAlaMetValTrpAlaLeuAsp 340
 DB 1050 ACCAAGGTGACGTATCTGAAGCAGAAGGAGCTGGGGGGGCCCATGGTCTGGGCACTGGAC 1109
 QY 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeu 360
 DB 1110 TTAGATGACTTTTGGCGGCTTCTCTGCAACCAAGGGCCGATACCCCTCATCCAGACGCTA 1169
 QY 361 ArgGlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLys 380
 DB 1170 CGGAGGAACTGAGTCTTCCATCTTGGCTTCAGGACCCAGAGCTTGAAGTTCCANA 1229
 QY 381 ProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGln 400
 DB 1230 CCAGGTTCAGCCCTCTGAACCTGAGCATGGCCCGCCCTGGACAAGACACGCTTCTGCCAG 1289
 QY 401 GlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyrSerCysAla 420
 DB 1290 GGCAAGCTGATGGGCTCTATCCCAATCTCTGGGAACGGTCCAGCTTCTACAGCTGTGCA 1349
 QY 421 AlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerAsnSerCysLys 440
 DB 1350 GGGGGGGGCTGTTCAGCAAGCTGCCCGACAGGCTGGTGTTCAGCACTCTCTGCAAA 1409
 QY 441 CysCysThrTrpAsn 445
 DB 1410 TGCTGCACCTGGAAT 1424

RESULT 7
 AAZ21848
 ID AAZ21848 standard; DNA; 1656 BP.
 XX
 AC AAZ21848;
 XX 10-DEC-1999 (first entry)
 DT
 XX
 DE MO-13B clone of human Chitinase, with noncoding 5'/3' regions.
 XX chitin; fungal infection; immunocompromised; AIDS; chemotherapy;
 KW organ transplant; parasite; chitin-binding; allele; vector;
 KW truncated protein; ds.
 XX
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 27..1427
 FT /tag= a
 FT /product= Human_Chitinase
 FT 27..89
 FT /tag= b
 FT /note= "Signal peptide"
 FT 90..1427
 FT /tag= c
 FT /note= "Mature peptide"
 XX
 PN WO9946390-A1.

XX 16-SEP-1999.
 XX 12-MAR-1999; 99WO-US05343.
 XX 12-MAR-1998; 98US-0039198.
 XX (ICOS-) ICOS CORP.
 XX Gray PW, Tjoelker LW;
 XX WPI; 1999-551417/46.
 XX P-PSDB; AAY42426.
 XX Novel chitin-binding fragments of human chitinase used to treat fungal
 XX infections in animals
 XX Example 1; Page 59-62; 83pp; English.
 XX This is the nucleotide sequence of an allelic form of the human
 XX chitinase enzyme, which is capable of degrading Chitin (a linear
 XX homo polymer of beta-1,4-linked N-acetylglucosamine residues).
 XX Chitinase fragments can be used to screen for proteins or other
 XX molecules that specifically bind to the chitin-binding domain of human
 XX chitinase or that modulate its activity. These compounds are useful for
 XX immunization, as well as for purifying chitinase, as well as for
 XX detection and quantification of chitinase. Polynucleotide fragments of
 XX the invention are useful as a source of probes and primers, and to
 XX express the proteins recombinantly. The chitinase fragments, when
 XX conjugated to antifungal compounds, are used to treat animals,
 XX especially humans, infected with chitin-containing parasites such as
 XX fungi. Fungal infection treated include candidiasis, aspergillosis,
 XX coccidioidomycosis, blastomycosis, paracoccidioidomycosis,
 XX mucormycosis, histoplasmosis, cryptococcosis, chromoblastomycosis,
 XX sporotrichosis, and dermatophycoses.
 XX Chitin can be degraded by the enzyme chitinase. Use of whole chitinase
 XX protein for treating infections, especially fungal infections, is
 XX problematic. In view of the increasing incidents of life-threatening
 XX fungal infection in e.g. immunocompromised individuals, there exists a
 XX need for identifying new compounds for treating fungal infection. The
 XX chitin-binding fragments of the present invention provide this need.

XX SQ Sequence 1656 BP; 365 A; 497 C; 447 G; 347 T; 0 other;

Alignment Scores:
 Pred. No.: 2, 88e-200 Length: 1656
 Score: 2392.00 Matches: 444
 Percent Similarity: 99.78% Conservative: 0
 Best Local Similarity: 99.78% Mismatches: 1
 Query Match: 99.75% Indels: 0
 DB: 20 Gaps: 0

US-10-004-219B-10 (1-445) x AA221848 (1-1656)

Qy 1 AlaLysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyAlaLahargPhe 20
 Db 90 GCAAACTGGTCTGCTACTTCCAACTGGGCCAGTACAGACAGGGGAGGCTCGCTC 149
 Qy 21 LeuProLysAspLeuAspProSerLeuCysThrHisLeuLeuTyrAlaPheAlaGlyMet 40
 Db 150 CYGCCCAAGGACTTGGACCCCAAGCCCTTTGCACCCACCTCATCTAGCCCTTCGCTGGCATG 209
 Qy 41 ThrAsnHisGlnLeuSerThrThrGlnTrpAsnAspGluThrLeuTyrGlnGluPheAsn 60
 Db 210 ACCAACCCACCTGAGCACCCTGAGTGGGAATGACGAGACTCTTACCAAGAGTTCAAT 269
 Qy 61 GlyLeuLysLysMetAsnProLysLeuLysThrLeuLeuAlaileGlyLysThrAsnPhe 80
 Db 270 GGCCTGAAGAAGATGAATCCCAAGCTGAGACCCCTGTTAGCCATCGGAGGCTGGAATTC 329
 Qy 81 GlyThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsn 100
 Db 330 AGCACTCAGAAGTTCACAGATATGTTAGTCCAGCGCCCAACCAACCGTCAGACCTTTGTCAAC 389

Qy 101 SerAlaileArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyr 120
 Db 390 TCGCCATCAGCTTCTGGGCAATATACAGCTTTCAGCGCCTTGACCTTGAGCTGGGAGTAC 449
 Qy 121 ProGlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAsp 140
 Db 450 CCAGGAAGCCAGGGAGGCGCTTCGCTAGACAGGAGCGCTTTCACAAACCTGGTGCAGGAC 509
 Qy 141 LeuAlaAsnAlaPheGlnGlnGlnAlaGlnThrSerGlyLysGluArgLeuLeuSer 160
 Db 510 TTGCCAATGCTTCCAGCAGGAAGCCAGACCTTCAGGAAGGAACGCTTCTTCTGAGT 569
 Qy 161 AlaAlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAla 180
 Db 570 GCAGGGTTCAGCTGGGCAGACCTATGTGTGATCTGTGATACGAGGTGCACAAATGCC 629
 Qy 181 GlnAsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysVal 200
 Db 630 CAGAACCTGGATTTTGTCAACCTTATGGCCTTACGACTTCCATGGCTTTGGGAGAAGTTC 689
 Qy 201 ThrGlyHisAsnSerProLeuTyrLysArgGlnGluSerGlyValAlaAsnSerLeu 220
 Db 690 ACGGACATACAGCCCTCTTACAAGAGGCAAGAGAGTGGTGCAGCAGCCAGCTTC 749
 Qy 221 AsnValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeu 240
 Db 750 AACGTGATGCTGTGTGCACAGCTGGCTGCAGAGGGGAGCCCTGCCAGCAAGCTGATC 809
 Qy 241 LeuGlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgVal 260
 Db 810 CTTGGCATGCTTACCTACGGAGCGCTCTTCACTTGGCTCTCATCAGACACCAAGAGTG 869
 Qy 261 GlyAlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyMetLeu 280
 Db 870 GGGGCCCCAGCCAGCAGGGTCTGGCACTCCAGGCCCTTCCACAGGAGGAGGGATGCTG 929
 Qy 281 AlaTyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLys 300
 Db 930 GCCTACTATGAAGTCTGCTCTCGAAGGGGGCCCAACAGAGAAATCCAGGATCAGAAG 989
 Qy 301 ValProTyrIlePheArgAspAsnGlnTrpValGlyPheAspValGluSerPheLys 320
 Db 990 GTGCCCTACATCTTCGGGCAACACAGTGGGCTTTGATGATGTGGAGAGCTTCAAA 1049
 Qy 321 ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyAlaMetValTrpAlaLeuAsp 340
 Db 1050 ACCAAGGTGACGTATCTGAAGCAGAGGGGACTGGGGCGGGCCATGGTCTGGGCATGGAC 1109
 Qy 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeu 360
 Db 1110 TTAGATGACTTTGGCGGCTTCTCTGCAACAGGGCCGATACCCCTCATCCAGAGCTA 1169
 Qy 361 ArgGlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLys 380
 Db 1170 CGCAGGAAGTCTGAGTCTTCCATCTTGGCTTTCAGGACCCCGAGAGCTTGAAGTTCAGAAA 1229
 Qy 381 ProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGln 400
 Db 1230 CCAGGTGACGCTCTGAACCTGAGATGGCCCGCCAGCTGGCAAGACACGCTTCTGGCAG 1289
 Qy 401 GlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyrSerCysAla 420
 Db 1290 GGCAAGTCTGATGGCTCTATCCCAATCTCTGGGACCGTCCAGCTTCTACAGCTGTGCA 1349
 Qy 421 AlaGlyArgLeuPheGlnLysSerCysProThrGlyLeuValPheSerAsnSerCysLys 440
 Db 1350 GCGGGGGGCTGTTCCAGAAAGCTGCCGACAGCGCTGGTGTTCAGCAACTCTCTGCAAA 1409
 Qy 441 CysCysThrTrpAsn 445
 Db 1410 TGCTGCACCTGGAAT 1424

RESULT 8

AAD03760
 ID AAD03760 standard; cDNA; 1656 BP.
 AC AAD03760;
 XX
 DT 19-JUN-2001 (first entry)
 XX Human chitinase cDNA from clone pMO-13B.
 DE
 XX Human; antifungal; chitinase; immunoglobulin; Ig; therapy;
 KW fungal infection; candidiasis; aspergillosis; coccidioidomycosis;
 KW blastomycosis; paracoccidioidomycosis; histoplasmosis; cryptococcosis;
 KW chromoblastomycosis; sporotrichosis; mucormycosis; dermatophytoses;
 KW clone pMO-13B; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 27..1427
 FT /tag= a
 FT /product= "Human chitinase from clone pMO-13B"
 FT sig_peptide 27..89
 FT /tag= b
 FT mat_peptide 90..1424
 FT /tag= c
 FT /product= "Human mature chitinase from clone pMO-13B"
 XX WO200123430-A2.
 PN
 XX
 PD 05-APR-2001.
 XX
 PF 28-SEP-2000; 2000WO-US26960.
 XX
 PR 30-SEP-1999; 99US-0409918.
 XX
 XX (ICOS-) ICOS CORP.
 PA
 XX
 XX Allison DS, Dietsch GN, Gray PW, Shaw KD, Steiner BH;
 PI
 XX WPI; 2001-266141/27.
 DR P-PSDB; AAE00433.
 XX
 XX Novel chitinase immunoglobulin fusion product, useful for treating
 PT fungal infections and reducing the amount of a non-chitinase antifungal
 PT agent needed for the treatment
 XX
 PS Claim 2; Page 34-36; 39pp; English.
 XX
 CC The present invention relates to a chitinase immunoglobulin (Ig) fusion
 CC product, comprising a human chitinase fused to at least a portion of an
 CC immunoglobulin chain. The fusion product is useful for treating fungal
 CC infections (mycoses) such as candidiasis, aspergillosis, blastomycosis,
 CC coccidioidomycosis, paracoccidioidomycosis, histoplasmosis, mucormycosis,
 CC cryptococcosis, chromoblastomycosis, sporotrichosis and dermatophytoses.
 CC The fusion protein is useful for reducing the amount of non-chitinase
 CC antifungal agent needed to exert an antifungal activity. The fusion
 CC protein is also useful for preparing a medicament for the prophylactic
 CC or therapeutic treatment of fungal infections. Chitinase immunoglobulin
 CC fusion product has unexpectedly improved serum half-life and formulation
 CC properties. The present sequence is human chitinase cDNA from clone
 CC pMO-13B. Chitinase enzyme degrades chitin which is a homopolymer of
 CC beta-(1,4)-linked N-acetylglucosamine residues.
 XX
 SQ Sequence 1656 BP; 365 A; 497 C; 447 G; 347 T; 0 other;
 XX

Alignment Scores:

Pred. No.: 2,88e-200 Length: 1656
 Score: 2392.00 Matches: 444
 Percent Similarity: 99.78% Conservative: 0
 Best Local Similarity: 99.78% Mismatches: 1
 Query Match: 99.75% Indels: 0
 DB: 22 Gaps: 0

US-10-004-219B-10 (1-445) x AAD03760 (1-1656)
 QY 1 AlaLysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPhe 20
 DB 90 GCAAAACTGGTCTGCTACTTACCAACTGGGCCAGTACAGACAGGGGAGGCTCGCTTC 149
 QY 21 LeuProLysAspLeuAspProSerLeuCysThrHisLeuLeuTyrAlaPheAlaGlyMet 40
 DB 150 CTGCCCAAGGACTTGGACCCCGAGCCCTTTGCACCCACCTCATCTACGCTTCGCTGGCATG 209
 QY 41 ThrAsnHisGlnLeuSerThrThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsn 60
 DB 210 ACCAAACCCAGCTCAGCACCACCTGAGTGAATGACGAGACTCTCTACGAGAGTTCAAT 269
 QY 61 GlyLeuLysLeuMetAsnProLysLeuLysThrLeuLeuAlaLeuGlyGlyTrpAsnPhe 80
 DB 270 GGCCTGAAGAAGATGAATCCCAAGCTGAAGCCCTTTAGCCCTCGAGGCTGGGAATTC 329
 QY 81 GlyThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsn 100
 DB 330 AGCACTCAGAAGTTTACAGATATGTCAGCCAGCGCAACCAACCGTCAGACCTTTGTCAAC 389
 QY 101 SerAlaLeuArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyr 120
 DB 390 TCGGGCATCAGGTTTCTCGCAAAATACAGCTTTGACGGCCTTGACCTTGACCTGGGAGTAC 449
 QY 121 ProGlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAsp 140
 DB 450 CCAGGAAGCAGGGAGGCCCTCGCTGATAGCAAGAGCGCTTCACAACTGGTGTGAGGAC 509
 QY 141 LeuAlaAsnAlaPheGlnGlnGlnAlaGlnThrSerGlyLysGluArgLeuLeuSer 160
 DB 510 TTGGCCCAATGCCCTTCCAGCAGGAGAGCCAGACCTCAGGGAGAGAAACGCTTCTCTGAGT 569
 QY 161 AlaAlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysLeuAla 180
 DB 570 GCAGCGGTTCCAGCTGGGCGAGACCTATGTGGATGTGGATAGCAGGTGGACAAAATCGCC 629
 QY 181 GlnAsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysVal 200
 DB 630 CAGAACTTGGATTTTGTCAACCTTATGGCTCAGACTTCCATGGCTCTTGGGAGAGGTC 689
 QY 201 ThrGlyHisAsnSerProLeuTyrLysArgGlnGluSerGlyAlaAlaAlaSerLeu 220
 DB 690 ACGGACATACAGACCCCTCTTACAAGAGGCAAGAGAGAGTGGTGCAGACCCAGCCCTC 749
 QY 221 AsnValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuLeu 240
 DB 750 AACGTGGATGCTGTGTGCAACAGTGGCTGCAGAAAGGGAGCCCTTGCAGCAAGCTGATC 809
 QY 241 LeuGlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgVal 260
 DB 810 CTTGGCATGCCCTACTACGACGCTTCTTCACTATGGGCTCTCTCATCAGACACCAAGATG 869
 QY 261 GlyAlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyMetLeu 280
 DB 870 GGGGGCCCGAGCCACAGGCTTGGCACTCCAGGCCCTTCCACCAAGGAGGAGGATGCTG 929
 QY 281 AlaTyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLys 300
 DB 930 GCCTACTATGAAGTCTGCTCTCTGGAAGGGGGGCCACCAACAGAGAATCCAGGATCAGAG 989
 QY 301 ValProTyrIlePheArgAspAsnGlnTrpValGlyPheAspAspValGluSerPheLys 320
 DB 990 GFGCCCTACATCTCCGGGACCAACAGTGGGTGGGCTTTGATGATGTGGAGAGCTTCAAA 1049
 QY 321 ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyGlyAlaMetValTrpAlaLeuAsp 340
 DB 1050 ACCAAGGTGAGTATCTGAAGCAGAAAGGAGCTGGGGCGGGCCATGGTCTGGGCACTGGAC 1109
 QY 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeu 360

Db 1110 TTAGATGACTTTGGCGGCTTCTCTGCAACAGGCGCGATACCCCTCATCCAGAGCTTA 1169
 Qy 361 AtgGlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLys 380
 Db 1170 CGGAGGAACTGAGCTTCATCTGCTTTCAGGACCCAGCTTGAAGTTCAAAA 1229
 Qy 381 ProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGln 400
 Db 1230 CCAGGTGAGCCCTGCAACCTGAGCATGCGCCAGCCCTGGACAGACACGTTCTGCCAG 1289
 Qy 401 GlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyrSerCysAla 420
 Db 1290 GGCAAGCTGATGGGCTCTATCCCAATCTCCGGAACGGTCCAGCTTCACAGCTGTGCA 1349
 Qy 421 AlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerAsnSerCysLys 440
 Db 1350 GGGGGGGGCTGTTCCAGCAAGAGCTGCCGACAGGCTGGTGTTCAGCAACTCTTCGCAA 1409
 Qy 441 CysCysThrTrpAsn 445
 Db 1410 TGCTGCACCTGGAAT 1424

RESULT 9

ABL57381
 ID ABL57381 standard; cDNA; 1656 BP.

AC ABL57381;

DT 12-AUG-2002 (first entry)

XX Human chitinase cDNA clone MO-13B.

DE Chitinase; enzyme; human; fungicide; antifungal; infection;
 KW candidiasis; aspergillosis; coccidioidomycosis; blastomycosis;
 KW paracoccidioidomycosis; histoplasmosis; cryptococcosis;
 KW Chromoblastomycosis; sporotrichosis; mucormycosis; dermatophytosis;
 KW Pneumocystis; gene; ss.
 OS Homo sapiens.

XX Key Location/Qualifiers
 FH 27..1427
 CDS /tag= a

FT sig_peptide /product= "Chitinase"

FT 27..89 /tag= b

FT mat_peptide 90..1424

FT /tag= c

PN US6372212-B1.

XX 16-APR-2002.

PD 16-JUN-1997; 97US-0877599.

PF 14-JUN-1996; 96US-0663618.

PR (ICOS-) ICOS CORP.

XX Gray FW;

PI WPI; 2002-442449/47.

DR P-PSDB; ABB76292.

XX Co-administering chitinase to improve the effectiveness of fungicidal
 PT drugs e.g. amphotericin B or itraconazole, useful for treating fungal
 PT infections e.g. candidiasis, coccidioidomycosis and blastomycosis -
 XX Example 1; Column 31-34; 26pp; English.
 PS The present sequence is that of cDNA clone MO-13B (ATCC 98078)
 CC encoding human chitinase (see ABB76292). The clone was isolated
 CC from a cDNA library prepared from peripheral blood monocyte-derived

CC macrophages using a probe based on human chitinase clone MO-218
 CC (see ABL57380). The 2 clones contain a single nucleotide
 CC difference in the coding region, causing amino acid position 81 of
 CC the mature protein to be either serine or glycine. Northern blots
 CC showed highest chitinase gene expression in lung and ovary tissues.
 CC Expression in lung is consistent with a protective role against
 CC pathogenic organisms that contain chitin. The invention provides
 CC human chitinase polynucleotides and polypeptides, and materials and
 CC methods for the recombinant production of human chitinase products,
 CC which are expected to be useful as products for treating fungal
 CC infections or for the development of such products. Human
 CC chitinase has a synergistic effect on the actions of other
 CC fungicides. It can be administered to improve the antifungal
 CC activity of a non-chitinase antifungal agent, especially
 CC amphotericin B or itraconazole, in the treatment of a fungal
 CC infection such as candidiasis, aspergillosis, coccidioidomycosis,
 CC blastomycosis, paracoccidioidomycosis, histoplasmosis,
 CC cryptococcosis, chromoblastomycosis, sporotrichosis, mucormycosis,
 CC dermatophytoses and Pneumocystis infections (all claimed). In
 CC particular, the fungal infection involves Candida, Aspergillus
 CC and/or Cryptococcus spp., whose growth is not effectively
 CC inhibited by contact with human chitinase alone.

XX SQ Sequence 1656 BP; 365 A; 497 C; 447 G; 347 T; 0 other;

Alignment Scores:

Pred. No.: 2,886-200 Length: 1656
 Score: 2392.00 Matches: 444
 Percent Similarity: 99.78% Conservative: 0
 Best Local Similarity: 99.78% Mismatches: 1
 Query Match: 99.75% Indels: 0
 DB: 24 Gaps: 0

US-10-004-219B-10 (1-445) x ABL57381 (1-1656)

Qy 1 AlalysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPhe 20
 Db 90 GCATAAACTGGTCTGCTACTTCCACCACTGGGCCAGTAGTACAGACAGGGGAGGCTCGCTTC 149
 Qy 21 LeuProLysAspLeuAspProSerLeuCysThrHisLeuLeuTyrAlaPheAlaGlyMet 40
 Db 150 CTGCCCAAGGACTTTGGACCCAGCCCTTTGCACCCACCTCATCTACGCCCTTCCTGGCATG 209
 Qy 41 ThrAsnHisGlnLeuSerThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsn 60
 Db 210 ACCAACCCAGCTGAGCACCACCTGAGTGGATGACGAGACTCTCTACGAGAGTTCAAT 269
 Qy 61 GlyLeuLysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPhe 80
 Db 270 GGCCTGAAGAGATGAATCCCAAGCTGAAGACCTGTTAGCCATCGAGGCTGGAATTC 329
 Qy 81 GlyThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsn 100
 Db 330 AGCACTCAGAAAGTTCCACAGATATGTTAGCCAGGCCCAACACCTTCAGACCTTTGTCAAC 389
 Qy 101 SerAlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyr 120
 Db 390 TCGGCCATCAGGTTTCTGGGCAATACAGCTTTGAGGCTTGACCTTGACTGGGAGTAC 449
 Qy 121 ProGlySerGlnGlySerProAlaValAspLysGluArgPheThrLeuValGlnAsp 140
 Db 450 CCAGGAAGCCAGGGGAGCCCTGCGCTAGACAGGAGCGCTTCACAAACCTGTCAGGAGAC 509
 Qy 141 LeuAlaAsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuSer 160
 Db 510 TTGGCAATATGCTTCCAGCAGGAGGCCAGACTCCAGGAGAGAACCGCTTCTCTGAGT 569
 Qy 161 AlaAlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAla 180
 Db 570 GCAGCGGTTCCAGCTGGGCGAGACCTATGTGGATGCTGGATAGAGGTGACAAATCGCC 629
 Qy 181 GlnAsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysVal 200

```
Db 630 CAGAACCTGGATTTCACACCTTATGGCTACGACTTCATGGCTCTTGGGAGAAAGGTC 689
QY 201 ThrGlyHisAsnSerProLeuTyrLysArgGlnGluSerGlyAlaAlaSerLeu 220
   |||||
Db 690 ACGGACATACAGCCCTCTACAGAGGCAAGAGAGTGGTCGACGACGCCCTTC 749
QY 221 AsnValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuIle 240
   |||||
Db 750 AACGTGGATGCTGCTGTGCAACAGTGGCTGCAGAGGGACCCCTGCCAGCAAGCTGATC 809
QY 241 LeuGlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgVal 260
   |||||
Db 810 CTTGGCATGCTACCTACCGAGCCTCTTCACACTGGCTCTTCATCAGACACCAAGTG 869
QY 261 GlyAlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyMetLeu 280
   |||||
Db 870 GGGCCCCCAGCCACAGGGTCTGCACTCCAGGCCCTTCACCAAGAGAGGAGGATGCTG 929
QY 281 AlaTyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLys 300
   |||||
Db 930 GCCTACTATGAAGTCTGCTCTGGAAGGGGGCCCAACCAACAGAGAATCCAGGATCAGAAG 989
QY 301 valProTyrIlePheArgAspGlnTrpValGlyPheAspValGluSerPheLys 320
   |||||
Db 990 GTGCCCTATACATCTCCGGGACACCAAGTGGTGGCTTTGATGATGTGAGAGCTTCAAA 1049
QY 321 ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyValAlaMetValTrpAlaLeuAsp 340
   |||||
Db 1050 ACCNAGGTAGCTATCTGAAGCAGAGGAGACTGGGGGGGCCCATGCTCTGGGCACTGGAC 1109
QY 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeu 360
   |||||
Db 1110 TTAGATGACTTTGCGGCTCTCTGCAACCAAGGGCCGATACCCCTCATCCAGAGCTTA 1169
QY 361 ArgGlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLys 380
   |||||
Db 1170 CGGCAGGAATGAGTCTTCATACTTGCCTTCAGGCACCCCAAGAGCTTGAAGTTCCAAA 1229
QY 381 ProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGln 400
   |||||
Db 1230 CCAGGTACGCCCTCTGACCTGAGCATGGCCCGCCCTGGACACAGACAGCTTCTGCCAG 1289
QY 401 GlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyrSerCysAla 420
   |||||
Db 1290 GGCAGAGCTGATGGCTCTATCCCAATCTCGGAAACGGTCCAGCTTCTACAGCTGTGCA 1349
QY 421 AlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerIleSerCysLys 440
   |||||
Db 1350 GCGGGCGGCTGTTCCAGCAAGCTGCCGACAGGCTGGTGTTCAGCAACTCCTGCAAA 1409
QY 441 CysCysThrTrpAsn 445
Db 1410 TGCTGCACCTGGAT 1424

RESULT 10
AAT89181
ID AAT89181 standard; cDNA; 1768 BP.
XX
AC AAT89181;
XX
XX CysCysThrTrpAsn 445
DT 27-APR-1998 (first entry)
DE Human chitotriosidase variant cDNA.
XX
XX Chitotriosidase; tissue remodelling disorder; diagnosis; therapy;
KW rheumatoid arthritis; atherosclerosis; human; ss.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FT CDS 124..1524
FT /*tag= a
XX
```

```
PN W09736917-Al.
XX
PD 09-OCT-1997.
XX
PF 21-MAR-1997; 97WO-US05072.
XX
PR 29-MAR-1996; 96US-0014295.
XX (HUMA-) HUMAN GENOME SCI INC.
PA (SMIK) SMITHKLINE BEECHAM CORP.
XX
XX Chopra A, Gentz R, Kirkpatrick RB, Ni J, Thotakura NR;
PI P-PSDB; AAW31498.
DR WPI; 1997-503041/46.
XX
XX New isolated human chitotriosidase gene - used to develop products
PT for the diagnosis and treatment of tissue remodeling disorders, e.g.
PT rheumatoid arthritis
XX
PS Claim 1; Page 21-22; 34pp; English.
XX
CC This human cDNA sequence encodes chitotriosidase (see AAW31498). It
CC is a preferred variant of the chitotriosidase cDNA sequence given
CC in AAT89180. Also claimed are: (1) a method of diagnosing a tissue
CC remodelling disorder related to expression of a mutated
CC chitotriosidase protein in a host comprising carrying out nucleic
CC acid amplification; and (2) a method of detecting altered
CC expression of a chitotriosidase protein in a host comprising
CC contacting a bodily sample with an antibody. Chitotriosidase can
CC degrade extracellular matrix substrates with a similar carbohydrate
CC structure to chitin. It can be used to develop products which can
CC be used in the diagnosis and treatment of tissue remodelling
CC disorders such as rheumatoid arthritis or atherosclerosis.
CC Chitotriosidase nucleic acids can be used in the recombinant
CC production of the enzyme, and as probes or primers in diagnostic
CC assays.
XX
SQ Sequence 1768 BP; 401 A; 529 C; 474 G; 363 T; 1 other;

Alignment Scores:
Pred. No.: 3,84e-200 Length: 1768
Score: 2391.00 Matches: 444
Percent Similarity: 99.78% Conservative: 0
Best Local Similarity: 99.78% Mismatches: 1
Query Match: 99.71% Indels: 0
DB: 18 Gaps: 0

US-10-004-219B-10 (1-445) x AAT89181 (1-1768)
QY 1 AlaLysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyValAlaArgPhe 20
   |||||
Db 187 GCANAACTGGTCTGTCTACTTACCAACTGGGCCCACTACAGACAGGGGAGGCTCGTTC 246
QY 21 LeuProLysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet 40
   |||||
Db 247 CTGCCCAAGGACTTGACCCAGCCCTTGCACCCACCTCATCTACGCCCTTCGCTGCATG 306
QY 41 ThrAsnHisGlnLeuSerThrThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsn 60
   |||||
Db 307 ACCAACCCACGCTGAGCACCACCTGAGTGAATGACGAGACTCTCTACAGAGATTCAAT 366
QY 61 GlyLeuLysLysMetAsnProLysLysThrLeuLeuAlaIleGlyGlyTrpAsnPhe 80
   |||||
Db 367 GGCCTGAGAGAGATGAATCCCAAGCTGAAGCCCTTTAGCCATCGAGGCTGGATTC 426
QY 81 GlyThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsn 100
   |||||
Db 427 RGCACTCAGAAGTTCACAGATATGTCAGCCAGCGGCAACCAACCGTCAGACCTTTGTCAAC 486
QY 101 SerAlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyr 120
   |||||
Db 487 TCGGCCATCAGGTTTCTCGGCAATACAGCTTTGACGGCCTTGACCTTGAGTGGAGTAC 546
```

QY	121	ProGlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAsp	140
DB	547	CCAGGAAGCAGCGGGAGCCCTGCCGTAGACAAGAGCGCTTCAACCCCTGGTACAGCAC	606
QY	141	LeuAlaAsnAlaPheGlnGlnGlnAlaGlnThrSerGlyLysGluArgLeuLeuSer	160
DB	607	TTCGCCAATGCCTTCAGCAGGAAGCCAGACTCAGGGAAAGAACGCCCTTCTCTGAGT	666
QY	161	AlaAlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAla	180
DB	667	GCAGCGTTCCAGCTGGCAGACCTATGTGATGCTGGATACGAGGTGGACAAAATGCC	726
QY	181	GlnAsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysVal	200
DB	727	CAGAACCCTGATTGTTGTCAACCTTATGGCTTAGCACTTCCATGGCTCTTGGAAGGTC	786
QY	201	ThrGlyHisAsnSerProLeuTyrLysArgGlnGluSerGlyAlaAlaAlaSerLeu	220
DB	787	ACGGGACATAACAGCCCCCTCTCAAGAGGACAAAGAGAGTGGTGAGCAGCGCCCTC	846
QY	221	AsnValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuIle	240
DB	847	NAGCTGGATGCTGCTGTCAACAGTGGTGCAGAGGGGACCCCTGCCACAGCTGATC	906
QY	241	LeuGlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgVal	260
DB	907	CTTGCGATGCTACCTACGAGCGCTCTTCACTGGCTCTCATCAGACACCCAGAGTG	966
QY	261	GlyAlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyMetLeu	280
DB	967	GGGGCCCAAGCCACAGGCTGTGCACTTCCAGGCCCTTCAACAGGAAGAGGAGTCTG	1026
QY	281	AlaTyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLys	300
DB	1027	GCTACTATGAAGTGTCTCTCGAAGGGGGCCACCAAACAGAGATCCAGGATCAGAAG	1086
QY	301	ValProTyrIlePheArgAspAsnGlnTrpValGlyPheAspValGluSerPheLys	320
DB	1087	GTGCCCTACATCTCCGGGACACCCAGTGGTGGCTTGTATGATGTGGAGAGCTTCAAA	1146
QY	321	ThrLysValSerTyrLeuLysGlnLysGlyValahetValTrpAlaLeuAsp	340
DB	1147	ACCAAGGTCAGCTATCTGAAGCAGAAGGGGACTGGCGGGGCCATGGTCTGGGCATGGAC	1206
QY	341	LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeu	360
DB	1207	TTAGATGACTTTCGGGCTTCTCTGCACACAGGGCCGATACCCCCTCATCCAGACGCTA	1266
QY	361	ArgGlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLys	380
DB	1267	CGGCAGGAACGTAGTCTTCCATACCTTGCCTTCAGGCACCCACAGAGCTTGAAGTTCCAAA	1326
QY	381	ProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGln	400
DB	1327	CCAGGTCAGCCCTCTGAACCTGAGCATGGGCCCCAGCCCTGGACACACGTTCTGCAG	1386
QY	401	GlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerPheTyrSerCysAla	420
DB	1387	GGCAAGCTGATGGCTCTATCCAAATCTCTGGGAAGGGTCCAGCTTCTACAGCTGTGCA	1446
QY	421	AlaGlyArgLeuPheGlnGlnInserCysPheThrGlyLeuValPheSerAsnSerCysLys	440
DB	1447	GCGGGGGGCTGTCCAGCAAAAGTGCCCGACAGGCCCTGGTGTTCAGCAACTCTCTGCAA	1506
QY	441	CysCysThrTrpAsn	445
DB	1507	TGCTGCACCTGGAAAT	1521

RESULT 11
 AAT89180
 ID AAT89180 standard; cDNA; 1637 BP.
 XX

AAT89180;
 27-APR-1998 (first entry)
 Human chitotriosidase cDNA.
 Chitotriosidase; tissue remodelling disorder; diagnosis; therapy;
 rheumatoid arthritis; atherosclerosis; human; ss.
 Homo sapiens.

Key	Location/Qualifiers
CDS	27..1421
variation	/*tag= a
	330
	/*tag= b
	/note= "nucleotide 330 is G, or an insertion of
	GGTACA, in some variants of the sequence"
variation	1253
	/*tag= c
	/note= "nucleotide 1253 is C in some variants of
	the sequence"
variation	1426
	/*tag= d
	/note= "nucleotide 1426 is deleted in some variants
	of the sequence"
variation	1442
	/*tag= e
	/note= "nucleotide 1442 is deleted in some variants
	of the sequence"
variation	1495
	/*tag= f
	/note= "nucleotide 1495 is deleted in some variants
	of the sequence"
variation	1632
	/*tag= g
	/note= "nucleotide 1632 is G in some variants of
	the sequence"
variation	1634
	/*tag= h
	/note= "nucleotide 1634 is G in some variants of
	the sequence"

WO9736917-A1.
 09-OCT-1997. 97WO-US05072.
 21-MAR-1997; 96US-0014295.
 29-MAR-1996; 96US-0014295.
 (HUMA-) HUMAN GENOME SCI INC.
 (SMIK) SMITHKLINE BEECHAM CORP.
 Chopra A, Gentz R, Kirkpatrick RB, Ni J, Thorakura NR;
 WPI; 1997-503041/46.
 P-PSDB; AAW31497.
 New isolated human chitotriosidase gene - used to develop products
 for the diagnosis and treatment of tissue remodeling disorders, e.g.
 rheumatoid arthritis
 Claim 1; Page 18-19; 34pp; English.
 This human cDNA sequence includes a coding region of
 chitotriosidase (see AAW31497). It is deposited as ATCC 69953.
 Variants of the sequence are provided, including the preferred
 sequence variant given in AAT89181. Also claimed are: (1) a method
 of diagnosing a tissue remodelling disorder related to expression
 of a mutated chitotriosidase protein in a host comprising carrying
 out nucleic acid amplification; and (2) a method of detecting
 altered expression of a chitotriosidase protein in a host

Db 804 CTGGCATGCCCTACCTACGGACGCTCCTTCACACTGGCCTCCTCATCAGACACAGAGTG 863

261	Qy	GlyAlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyGlyMetLeu	280
864	Db	GGGGCCCCCAGCCACAGGGTCTGGGCACTCCAGGCCCTTCACCAAGAGAGAGGGATGCTG	923
281	Qy	AlaTyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLys	300
924	Db	GCCTACTATGAAGTCTGCTCTCTGGAAGGGGGCCACCAACAGAGAATCCAGGATCAGAAG	983
301	Qy	ValProTyrIlePheArgAspAsnGlnTrpValGlyPheAspAspValGluSerPheLys	320
984	Db	GTGCGCTACATCTTCGCGGACCAACCACTGGGTGGGCTTTGATGATGTGGAGAGCTTCAAA	104
321	Qy	ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyValAlaMetValTrpAlaLeuAsp	340
1044	Db	ACCAAGGTCACTATCTGAACAGCAGAGGGACTGGGCGGGGCCATGGTCTGGGCACTGGAC	110
341	Qy	LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeu	360
1104	Db	TTAGATGACTTTGCGGCTTCTCTGCAACAGGGCCGATACCCCTCATCCAGACGCTA	116
361	Qy	ArgGlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLys	380
1164	Db	CGGCAAGAACTGAGTCTTCATACTTCGCTTCAGGCAACCCAGAGCTTGAAGTTCACAA	122
381	Qy	ProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGln	400
1224	Db	CCAGGTCAAGCCTCTGAACCTGAGCATGGTCCCAGGCCCTGGACAAGACACGTTCTGCCAG	128
401	Qy	GlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyrSerCysAla	420
1284	Db	GGCAAGCTGATGGGCTCTATCCCAATCCTCGGGAACGGTCCAGCTTCTACAGCTGTGCA	134
421	Qy	AlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerAsnSerCysLys	440
1344	Db	GGGGGGCGGCTGTTCACGAAGAAGTGGCCGACAGGGCTGGTGTTCAGCAACTCTCTGC	140
441	Qy	CysCysThrTrpAsn 445	
1404	Db	TGCTGCACCTGGAAT 1418	
RESULT 12			
AAT50834			
ID	AAT50834 standard; cDNA; 1713 BP.		
XX	AAT50834;		
DT	24-MAR-1997 (first entry)		
XX	Human chitinase cDNA clone chi.39.		
DE	XX		
XX	Chitinase; chitotriosidase; chitin; infectious disease;		
KW	gene therapy; vaccine; lysosomal lipidosi; Gaucher's disease;		
KW	leishmaniasis; sarcoidosis; x-linked adrenoleukodystrophy;		
KW	multiple sclerosis; drug delivery; cosmetics; food; ss.		
XX	XX		
OS	Homo sapiens.		
XX	XX		
Key	Location/Qualifiers		
FH	13..1176		
CDS	/*tag= a		
FT	13..75		
FT	/*tag= b		
FT	76..1173		
FT	/*tag= c		
XX	XX		
PN	WO9640940-A2.		
PD	19-DEC-1996.		
XX	XX		
PF	06-JUN-1996; 96WO-NL00225.		
XX	XX		
PR	07-JUN-1995; 95US-0486839.		
XX	XX		

(UNAM) UNIV AMSTERDAM.

Aerts JMFg;

WPI: 1997-118698/11.

P-PSDB; AAW08585.

New human chitinase - used to treat or prevent infection by
chitin-contg. pathogens, in diagnosis and as additives to cosmetics,
foods, implants etc.

Claim 2; Page 42-43; 58pp; English.

A cDNA clone (AA050834), designated chi.39, codes for a 39 kDa human
chitinase (AAW08584) that is stable to many proteases, active at pH 3-
8 and up to 50 deg, and stable in the circulation. Clones chi.39
and chi.50 (see also AA050833) were isolated from a human macrophage
cDNA library using as probe a partial clone obtd. using primers
(see also AA050835-36) based on a chitotriosidase purified from a
type 1 Gaucher disease patient. The 2 cDNA clones are the result
of alternative splicing of RNA. Chitinase nucleic acid can be used
for large-scale prodn. of recombinant human chitinases, or can be
incorporated into a gene therapy vector to treat or prevent
infection by chitin-contg. pathogens.

Sequence 1713 BP; 380 A; 503 C; 465 G; 365 T; 0 other;

Alignment Scores:

Pred. No.:	1.98e-196	Length:	1713
Score:	2348.50	Matches:	443
Percent Similarity:	94.66%	Conservative:	0
Best Local Similarity:	94.66%	Mismatches:	1
Query Match:	97.94%	Indels:	25
DB:	18	Gaps:	1

US-10-004-219B-10 (1-445) x AAT50834 (1-1713)

2 LysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPheLeu 21
79 AAACCTGGTCTGCTACTTACCAACTGGGCCAGTACAGACAGGGGGAGGCTCGCTTCCTG 138

22 ProLysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMetThr 41
139 CCCAAGGACTTGGACCCACCGCTTGGACCCACCTCATCTAGCGCTTCGCTGGCATGACC 198

42 AsnHisGlnLeuSerThrThrGlnTrpAsnAspGluThrLeuTyrGlnGluPheAsnGly 61
199 AACCCACCACTGAGCACCACTGAGTGAATGACGAGACTCTTACCAAGAGTTCAATGGC 258

62 LeuLysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTyrAsnPheGly 81
259 CTGAAGACATGATCCCAAGCTGAAGACCTGTAGCCATCGAGGCTGGAAATTCGGC 318

82 ThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsnSer 101
319 ACTCAAGATTTCACAGATATGGTAGCCACCGGCCAACACCGTCAGACCTTTGTCAACTCG 378

102 AlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyrPro 121
379 GCCATCAGGTTTCTGGCAATATACAGCTTGTAGCGCTTGTAGCTGGAGTACCCA 438

122 GlySerGlnGlySerProAlaValaAspLysGluArgPheThrThrLeuValGlnAspLeu 141
439 GGAAGCCAGGGAGCCCTCGGTAGACAGAGGCGCTTCACAACTCGGTGTCAGGACTTG 498

142 AlaAsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuSerAla 161
499 GCCAATGCTTCCAGCAGGAGCCAGACCTCAGGAAGAACCGCTTCTCTGAGTGCA 558

162 AlaValProAlaGlyGlnThrThrValAspAlaGlyTyrGluValAspLysIleAlaGln 181
559 GCGGTTCAGCTGGGCAGACCTATGTGGATGTGGATACGAGGTGGACAAATCGCCAG 618

182 AsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysValThr 201
619 AACCTGGATTTTGTCAACCTTATGSCCTACGACTTCCATGCTCTTTGGAGAAGCTCAGC 678

202 GlyHisAsnSerProLeuTyrLysArgGlnGluSerGlyValAlaAlaAsnSerLeuAsn 221
679 GGACATAACAGCCCTCTCAAGAGGCAAGAGAGTGTGTGACAGCAGCCAGCTCAAC 738

222 ValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuIleLeu 241
739 GTGGATGCTGTGTGCAACAGTGTGCAAGAGGGGACCCCTGCCAGCAGCTGATCTT 798

242 GlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgValGly 261
799 GGCATGCTTACTACGGGACGCTCTTTCACACTGGGCTCTCATCAGACACACAGAGTGGG 858

262 AlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyGlyMetLeuAla 281
859 GCCCCAGCCACAGGCTCTGGGCACTCCAGGCCCTTCCACCAAGGAAGAGGGATGCTGCC 918

282 TyrTyrGluValCysSerTyrLysGlyAlaThrLysGlnArgIleGlnAspGlnLysVal 301
919 TACTATGAAGTCTGCTCTGGAAGGGGCCACCAACAGAGNATCCAGGATCAGAAAGTG 978

302 ProTyrIlePheArgAspAsnGlnTrpValGlyPheAspAspValGluSerPheLysThr 321
979 CCTACATCTTCCGGGACCAACAGTGGGTGGCTTTGATGATGTGGAGAGCTTCAAAACC 1038

322 LysValSerTyrLeuLysGlnLysGlyLeuGlyGlyAlaMetValTrpAlaLeuAspLeu 341
1039 AAGGTCAGCTATCTGAAGCAGAGGAGTGGGCGGGGCTAGTGTCTGGGCACCTGGACTTA 1098

342 AspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeuArg 361
1099 GATGACTTTGGCGGCTTCTCTGCAACCGAGGCCATACCCCTCATCCAGACGCTACGG 1158

362 GlnGluLeuSer----- 365
1159 CAGGAACCTGAA-TGGGTAAAGCCTTAACTGCTGTGCATGTGAGGCCAGGTTTGCCTG 1217

366-----LeuProTyrLeuProSerGlyThrProGluLeuGlu 377
1218 TGGCAGCTGCTTTCAGCTGTAGTCTTCCATCTTCCCTTTCAGGCACCCAGAGCTTGA 1277

378 ValProLysProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThr 397
1278 GTTCCAAACACAGGTACAGCCCTCTGAACCTGAGCATGGCCCAAGCCCTTGGACACACG 1337

398 PheCysGlnGlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyr 417
1338 TTCTGCCAGGGCAAGCTGATGGGCTCTATCCCAATCTCGGGAACGGTCCAGCTTCTAC 1397

418 SerCysAlaAlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerAsn 437
1398 AGCTGTGACGGGGCGGCTGTTCAGCAAAAGTCTGCCGACAGGCGCTGTGTTCAGCAAC 1457

438 SerCysLysCysCysThrTrpAsn 445
1458 TCCTGCAAAATGTCGACCTGGAAT 1481

RESULT 13
AAH42013
ID AAH42013 standard; DNA; 1368 BP.
XX AC AAH42013;
XX DT 24-AUG-2001 (first entry)
XX DE Disease treatment related oligonucleotide SEQ ID NO: 3.
XX KW Disease treatment; infection; chronic occlusive pulmonary disease;
XX KW bronchial asthma; ds.

PN W0200222660-A2.
 XX 21-MAR-2002.
 XX 10-SEP-2001; 2001WO-US26015.
 XX 11-SEP-2000; 2000US-0659671.
 XX (HYSE-) HYSEQ INC.
 PA Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
 PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;
 XX WPI; 2002-292408/33.
 DR P-PSDB; ABB97580.
 XX An isolated polynucleotide for treating diseases associated with its
 PT encoded polypeptide such as cancer and multiple sclerosis -
 PT Claim 1; SEQ ID NO 404; 509pp; English.
 CC The present invention provides the protein and coding sequences of 444
 CC novel human proteins. These were isolated from expressed sequences tags
 CC (ESTs). They can be used to stimulate cell growth, to regulate
 CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
 CC e.g. in burn treatment, to regulate the immune system e.g. to treat
 CC multiple sclerosis, to regulate activin or inhibin e.g. to treat
 CC infertility, to regulate haemostasis or thrombolysis e.g. to treat
 CC stroke and cancer, to screen for drugs, to treat inflammatory conditions
 CC e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.
 CC Parkinson's disease. The present sequence is a coding sequence of the
 CC invention.
 XX SQ Sequence 1369 BP; 327 A; 393 C; 341 G; 308 T; 0 other;

Alignment Scores:
 Pred. No.: 1,36e-84 Length: 1369
 Score: 1072.50 Matches: 200
 Percent Similarity: 64.53% Conservative: 62
 Best Local Similarity: 49.26% Mismatches: 117
 Query Match: 44.72% Indels: 27
 DB: 24 Gaps: 4

US-10-004-219B-10 (1-445) x ABB59993 (1-1369)
 QY 56 TyrGlnGluPheAsnGlyLeuLysLysMetAsnProLysLeuLysThrLeuAlaIle 75
 DB 67 TATACGAATT-----GCAGCTCGGAACAGCCAGCTGAAACTCTCTGGCCATT 117
 QY 76 GlyGlyTyrAsnPheGlyThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArg 95
 DB 118 GGAGGCTGGAACTCGGACTGCCCCCTTCTACATGCCATGGTTCTACTCTTGAGACCGC 177
 QY 96 GlnThrPheValAsnSerAlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAsp 115
 DB 178 CAGACTTTCATCACTACCTCAGTCATCAAAATCTCTGGCCAGTATGAGTTTCAAGCGCTGGAC 237
 QY 116 LeuAspTyrGluTyrProGlySerGlnGlySerProAlaValAspLysGluArgPheThr 135
 DB 238 TTGTACTGGAGTACCTCGCTCTCTGGAGGCCCTCTCTCAGGACCAAGCATCTCTTCACT 297
 QY 136 ThrLeuValGlnAspLeuAlaAsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGlu 155
 DB 298 GTCTGTGTGAGGAATGCGTGAAGCTTTTGACAGAGGCCAAGCAGATCAACAGCCCC 357
 QY 156 ArgLeuLeuSerAlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGlu 175
 DB 358 AGGCTGATGTCAGTCTGTCAGTAGCTGCTGGCATCTCCAATATCCAGTCTGGCTATGAG 417
 QY 176 ValAspLysIleAlaGlnAsnLeuAspPheValAsnMetAlaTyrAspPheHisGly 195
 DB 418 ATCCCCCAACTGTACAGTACCTGGAGTACATCCATGTGATGACCTACGACCTCCATGCG 477

QY 196 SerTyrGluLysValThrGlyHisSerProLeuTyrLysArgGlnGluGluSerGly 215
 DB 478 TCCTGGAGGGCTACACTGGAGAGAACAGCCCTCTACAAATACCGACTGACACCGGC 537
 QY 216 AlaAlaAlaSerLeuAsnValAspAlaValGlnGlnTrpLeuGlnLysGlyThrPro 235
 DB 538 AGCAACGCTCACTCAATGTGATTATGTATGAATCTTGAAGGACAATGGAGCACCA 597
 QY 236 AlaSerLysLeuLeuLeuMetProThrTyrGlyArgSerPheThrLeuAlaSerSer 255
 DB 598 GCTGAGAGCTCATCTGTGGATTCTCTACCTATGACACAACTTCTCTGAGCAACCC 657
 QY 256 SerAspThrArgValGlyAlaProAlaThrGlySerGlyThrProGlyProPheThrLys 275
 DB 658 TCCAACTGGAATGGTGGCCCACTCTGTGTGCTCTGTCTGGCCCTATGCCAAG 717
 QY 276 GluGlyGlyMetLeuAlaTyrTrpGluValCysSerTrp-----LysGlyAlaThrLys 293
 DB 718 GAGTCTGGGATCTGGGCTTACTACGAGATCTGTACCTTCTCTGAAATAATGGAGCCACTCAG 777
 QY 294 GlnArgIleGlnAspGlnLysValProTyrIlePheArgAspAsnGlnTrpValGlyPhe 313
 DB 778 GGATGGATGCCCTCAGGAAGTGCCTTATGCTATCAGGCCAATGTGTGGTTGGCTAT 837
 QY 314 AspAspValGluSerPheLysThrLysValSerTyrLeuLysGlnLysGlyLeuGlyGly 333
 DB 838 GACAACTCAAGAGCTTCGATATTAAGGCTCAATGGCTTAAGCACACAAATTTGGAGGC 897
 QY 334 AlaMetValTrpAlaLeuAspLeuAspPheAlaGlyPheSerCysAsnGlnGlyArg 353
 DB 898 GCCATGGTCTGGGCCAATGATCTGGATCACTTCTGCAACAGGGCAG 957
 QY 354 TyrProLeuIleGlnThrLeuArgGlnGluLeuSerLeuProTyrLeuProSerGlyThr 373
 DB 958 TTTCCCTCAATCTCCACTGAAGAGCCCTC-----GGCCTG 996
 QY 374 ProGluLeuGluValProLysProGlyGlnProSerGluPro----- 387
 DB 997 CAGAGTGCAAGTTGCACGCGCTCCAGCTCAGCCCAATTGAGCCAATAACTGCTCCTCCAGT 1056
 QY 388 -----GluHisGlyProSerProGlyGlnAspThrPhe 398
 DB 1057 GGCACGGGAACGGGAGCGGAGTAGCAGCTCTGAGGAGGAGCTCGGAGGAGGAGTTC 1116
 QY 399 CysGlnGlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyrSer 418
 DB 1117 TGTGCTGTGAGAGCCCAAGGCTCTACCCCTGGGCAATAACAGAAATGCCTTCTGGCAC 1176
 QY 419 CysAlaAlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerAsnSer 438
 DB 1177 TGCGTGAATGGAGTCACTTACCAGCAGAACTGCCAGGCGCGGCTTGTCTTCGACACCAAGC 1236
 QY 439 CysLysCysCysThrTrp 444
 DB 1237 TGTGATTGCTCAACTGG 1254

Search completed: June 29, 2003, 21:27:02
 Job time : 266.223 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: June 29, 2003, 21:12:38 ; Search time 63.5244 Seconds
(without alignments)
2148.326 Million cell updates/sec

Title: US-10-004-219B-10
Perfect score: 2398
Sequence: 1 AKLVCFYTNWAGYRQGEARF.....QQSCPTGLVFNCKCTWN 445

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame_p2n.model -DEV=xlpl
-Q=/cgn2_1/USPTO.spool/US10004219/runat_24062003_160229_8184/app_query.fasta_1.1877
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-NODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10004219@cgn_1_1_108@runat_24062003_160229_8184 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPOP=6
-FCAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:
1: /cgn2_6/prodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/prodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/prodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/prodata/1/ina/PCTUS.COMB.seq:*
6: /cgn2_6/prodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2398	100.0	1636	4	US-09-039-198A-1
2	2398	100.0	1636	4	US-08-877-599-1
3	2398	100.0	1636	4	US-09-267-574-1
4	2394	99.8	1643	2	US-08-486-839-3
5	2394	99.8	1643	3	US-09-151-011-3
6	2394	99.8	1643	4	US-09-343-623-3
7	2392	99.7	1656	4	US-09-039-198A-3
8	2392	99.7	1656	4	US-08-877-599-3
9	2392	99.7	1656	4	US-09-267-574-3
10	2348.5	97.9	1713	2	US-08-486-839-5
11	2348.5	97.9	1713	3	US-09-151-011-5
12	2348.5	97.9	1713	4	US-09-343-623-5

13	1046.5	43.6	1433	1	US-08-694-915-1	Sequence 1, Appli
14	1046.5	43.6	1526	1	US-08-694-915-3	Sequence 3, Appli
15	1030	43.0	1681	5	PCT-US94-07754-4	Sequence 4, Appli
16	1025	42.7	1681	2	US-08-581-527-4	Sequence 4, Appli
17	824	34.4	2452	2	US-08-524-051-1	Sequence 1, Appli
18	803.5	33.5	1677	4	US-09-545-814-13	Sequence 13, Appli
19	803.5	33.5	1677	4	US-09-545-814-15	Sequence 15, Appli
20	803.5	33.5	1749	4	US-09-545-814-4	Sequence 4, Appli
21	803.5	33.5	1749	4	US-09-545-814-6	Sequence 6, Appli
22	803.5	33.5	2610	4	US-09-545-814-1	Sequence 1, Appli
23	803.5	33.5	2610	4	US-09-545-814-3	Sequence 3, Appli
24	797.5	33.3	1919	4	US-09-545-814-31	Sequence 31, Appli
25	797.5	33.3	1919	4	US-09-545-814-33	Sequence 33, Appli
26	762	31.8	1470	4	US-09-292-225-40	Sequence 40, Appli
27	762	31.8	1470	4	US-09-292-225-42	Sequence 42, Appli
28	762	31.8	1527	4	US-09-292-225-37	Sequence 37, Appli
29	762	31.8	1527	4	US-09-292-225-39	Sequence 39, Appli
30	762	31.8	1621	4	US-09-292-225-34	Sequence 34, Appli
31	762	31.8	1621	4	US-09-292-225-36	Sequence 36, Appli
32	738	30.8	1608	4	US-09-292-225-20	Sequence 20, Appli
33	738	30.8	1608	4	US-09-292-225-22	Sequence 22, Appli
34	738	30.8	1665	4	US-09-292-225-17	Sequence 17, Appli
35	738	30.8	1665	4	US-09-292-225-19	Sequence 19, Appli
36	738	30.8	1752	4	US-09-292-225-14	Sequence 14, Appli
37	738	30.8	1752	4	US-09-292-225-16	Sequence 16, Appli
38	660	27.5	1478	4	US-09-545-814-28	Sequence 28, Appli
39	660	27.5	1478	4	US-09-545-814-30	Sequence 30, Appli
40	499.5	20.8	1167	1	US-07-939-501A-6	Sequence 6, Appli
41	499.5	20.8	1167	1	US-08-448-398-10	Sequence 10, Appli
42	499.5	20.8	1320	1	US-07-939-501A-14	Sequence 14, Appli
43	499.5	20.8	1364	1	US-07-939-501A-13	Sequence 13, Appli
44	499.5	20.8	1405	1	US-07-939-501A-9	Sequence 9, Appli
45	486	20.3	1554	1	US-08-045-269C-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-039-198A-1
; Sequence 1, Application US/09039198A
; Patent No. 6200951
; GENERAL INFORMATION:
; APPLICANT: Tjoelker, Larry W.
; TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; STREET: 233 South Wacker Drive/6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/09/039,198A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 27866/34391
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1636 base pairs
; TYPE: nucleic acid

STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: CDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 2..1399
 FEATURE:
 NAME/KEY: mat peptide
 LOCATION: 65..1399
 US-09-039-198A-1

Alignment Scores:

Pred. No.: 3,596-238 Length: 1636
 Score: 2398.00 Matches: 445
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 4 Gaps: 0

US-10-004-219B-10 (1-445) x US-09-039-198A-1 (1-1636)

```

QY 1 AlaIysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgInGlnGlyGluAlaArgPhe 20
DB 65 GCAAACTGGTCTGCTACTTACCAACTGGGCCAGTACAGAGGGGGAGGCTCGCTTC 124
QY 21 LeuProLysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet 40
DB 125 CTGCCCNAGGACTTGGACCCAGCCCTTGGACCCACTCATCTACGCCCTCGCTGGCATG 184
QY 41 ThrAsnHisGlnLeuSerThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsn 60
DB 185 ACCAACCCAGCTGAGCACCCTGAGTGGAAATGACGAGACTCTCTACAGGAGTTCAAT 244
QY 61 GlyLeuLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyTyrAsnPhe 80
DB 245 GGCTTGAAGAAGATGAATCCCAAGCTGAAGACCTGTGTAGCCATCGAGGCTGGAAATTC 304
QY 81 GlyThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsn 100
DB 305 GGCACTCAGAGTTTACAGATATGGTAGCCAGCCCNACCAACCGTCAGACCTTTGTTCAC 364
QY 101 SerAlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyr 120
DB 365 TCGGCCATCAGTTTCTGCGCAAAATACAGCTTTGACGGCTTGACCTTGACTGGAGTAC 424
QY 121 ProGlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAsp 140
DB 425 CCAGGAAGCCAGGGAGCCCTGCGGTAGACAGGAGCGCTTCACACCCCTGGTACAGGAC 484
QY 141 LeuAlaAsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuLeuSer 160
DB 485 TTGGCCAAATGCCCTTCAGCAGGAGGCCAGACCTCAGGAGGAGGAGGAGGAGGAGGAGG 544
QY 161 AlaAlaValProAlaGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAla 180
DB 545 GCAGCGGTTCCAGCTGGCAGAGCCTATGTGATGCTGGATACGAGGTGGACAAATCGCC 604
QY 181 GlnAsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysVal 200
DB 605 CAGAACCTGGATTGTTCACACCTTATGGCTACGACTTCATCGCTTGGGGAGAGGTC 664
QY 201 ThrGlyHisAsnSerProLeuTyrLysArgGlnGluSerGlyAlaAlaLeuSerLeu 220
DB 665 ACGGGACATACAGCCCTCTACAGAGGCGCAGAGAGAGAGTGGTGGCAGCAGCCCTC 724
QY 221 AsnValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProLaserLysLeuIle 240
DB 725 AACGTGGATGCTGCTGTGCAACAGTGGCTGCAAGGGAGCCCTGCCAGCAGCTGATC 784
QY 241 LeuGlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgVal 260
DB 785 CTTGGCATGCTTACCTACGAGCAGCTCTTACACTGGCCCTCTCATCAGACACAGAGTG 844

```

```

QY 261 GlyAlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyGlyMetLeu 280
DB 845 GGGGCCCCAGCCACAGGCTCTGCACCTCCAGGCCCTTCAACAGGAGGAGGATGCTG 904
QY 281 AlaTyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLys 300
DB 905 GCCTACTATGAAGTCTGCTCTCTGGAAGGGGGCCACCAACAGAGAAATCCAGGATCAGAAG 964
QY 301 ValProTyrIlePheArgAspAsnGlnTrpValGlyPheAspValGluSerPheLys 320
DB 965 GTGCCCTCATCTTCCGGGACCAACAGTGGGTGGGCTTTGATGATGGAGAGCTTCAAA 1024
QY 321 ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyGlyValaMetValTrpAlaLeuAsp 340
DB 1025 ACCAAGTTCAGCTATCTGAAGCAGAGGAGCTGGGGGGGCCATGCTCTGGGCACTGGAC 1084
QY 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeu 360
DB 1085 TTAGATGACTTTGCCGGCTTCTCTGCAACAGGCGCATACCCCTCATCCAGACGCTA 1144
QY 361 ArgGlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLys 380
DB 1145 CGCAGGAACTCAGTCTTCCATACCTTTCCTTTCAGGCACCCAGAGCTTGAAGTTCCAAA 1204
QY 381 ProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGln 400
DB 1205 CCAGGTCAGCCCTCTGAACCTGAGCATGGCCCGCCCTGGACAGACACGTTCTGCCAG 1264
QY 401 GlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyrSerCysAla 420
DB 1265 GGCMAAGCTGATGGGCTCTATCCCAATCTCTCGGAACGGTCCAGCTTCTACAGCTGTGCA 1324
QY 421 AlaGlyArgLeuPheGlnSerCysProThrGlyLeuValPheSerAsnSerCysLys 440
DB 1325 GCGGGGCGGCTGTTCAGCAAGGCTGCCCGACAGGCTGGTGTTCAGCAACTCTCTGCAAA 1384
QY 441 CysCysThrTrpAsn 445
DB 1385 TGCTGCACCTCGAAT 1399

```

RESULT 2
 US-08-877-599-1
 ; Sequence 1, Application US/08877599
 ; Patent No. 6372212
 ; GENERAL INFORMATION:
 ; APPLICANT: Gray, Patrick W.
 ; TITLE OF INVENTION: Chitinase Materials and Methods
 ; NUMBER OF SEQUENCES: 17
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSES: Marshall, O'Toole, Gerstein, Murray & Borun
 ; STREET: 6300 Sears Tower, 233 South Wacker Drive
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: United States of America
 ; ZIP: 60606-6402
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/877,599
 ; FILING DATE:
 ; CLASSIFICATION: 514
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/663,618
 ; FILING DATE: 14-JUN-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Rin-Laures, Li-Hsien
 ; REGISTRATION NUMBER: 33,547
 ; REFERENCE/DOCKET NUMBER: 278666/33994
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 312/474-6300

TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1636 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 2..1399
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 65..1399
US-08-877-599-1

Alignment Scores:

Pred. No.: 3,598-238 Length: 1636
Score: 2398.00 Matches: 445
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-10-004-219B-10 (1-445) x US-08-877-599-1 (1-1636)

QY 1 AlaLysLeuValCysThrPheThrAsnTrpAlaGlnTyrArgGlnGlyClnuAlaArgPhe 20
DB 65 GCAAACTGGTCTGCTACTTCCAACTGGGCCCATGACAGAGGGGAGGCTCGCTTC 124
QY 21 LeuProLysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet 40
DB 125 CTGCCAAGGACTTGGACCCAGCCCTTGCACCCACCTCATCTAGCCCTTCCTGGCATG 184
QY 41 ThrAsnHisGlnLeuSerThrThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsn 60
DB 185 ACCAACCCAGCTGAGCACCACCTGAGTGGATGGAATGACGAGACTCTCTACAGGAGTTCAAT 244
QY 61 GlyLeuLysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyClyTrpAsnPhe 80
DB 245 GGCCTGAAGAAGATGAATCCCAAGCTGAAGACCCCTGTGTAGCCATCGAGGCTGGAAATTC 304
QY 81 GlyThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsn 100
DB 305 GGCACCTCAGAGTTTCACAGATATGTAGACCGCCGACGACGACGACGACGACGACGACGAC 364
QY 101 SerAlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyr 120
DB 365 TCGGCCCATCAGGTTCTCGCGCAAAATACAGCTTTGACGGCCTTGACCTTGAGTGGAGTAC 424
QY 121 ProGlySerGlnGlySerProAlaValAlaAspLysGluArgPheThrThrLeuValGlnAsp 140
DB 425 CCAGGAAGCCAGGGAGCCCTCGCGTAGACAAAGGAGCGCTTCACACCCCTGGTACAGGAC 484
QY 141 LeuAlaAsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuLeuSer 160
DB 485 TTGGCCCATGCTTCCAGCAGGAGCCCAAGACTCAGGAAGAGACCCCTTCTCTGAGT 544
QY 161 AlaAlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAla 180
DB 545 GCAGCGGTTCCAGCTGGGCAGACCTATGTGGATGCTGGATACGAGGTGGACAAATCGCC 604
QY 181 GlnAsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysVal 200
DB 605 CAGAACCTGGATTTTGTCAACCTTATGGCTACGACTTCATGGCTCTTTGGGAGAGGTC 664
QY 201 ThrGlyHisAsnSerProLeuTyrLysArgGlnGluLysSerGlyAlaAlaAsnLeu 220
DB 665 ACGGGACATAACAGCCCTCTTCAAGAGCGCAGAGAGAGTGTGCAGACCCAGCCCTC 724
QY 221 AsnValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuIle 240

DB 725 AACGTGATGCTGCTGTGCAACAGTGGCTGCAGAAAGGGGCCCTTGCAGCAAGCTGATC 784
QY 241 LeuGlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgVal 260
DB 785 CTGGCATGCTTACCTACGAGCGCTCTTCACTTACCTGGCTCTCTATCAGACACCAAGAGTG 844
QY 261 GlyAlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyGlyMetLeu 280
DB 845 GGGGCCCCAGCCACAGGCTTGGCACTCCAGGCCCTTCCACCAAGGAAGGAGGATGCTG 904
QY 281 AlaTyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLys 300
DB 905 GCCTACTATGAAGTCTGCTCTCGGAAGGGGGCCCAACCAACAGAGAATCCAGGATCAGAAG 964
QY 301 ValProTyrIlePheArgAspAsnGlnTrpValGlyPheAspAspValGluSerPheLys 320
DB 965 GTGCCCTACATCTTCCGGGACACACGATGGGTGGGCTTTGATGTGGAGAGCTTCAAA 1024
QY 321 ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyGlyAlaMetValTrpAlaLeuAsp 340
DB 1025 ACCAAGGTCAGCTATCTGAAGCAGAAAGGACTGGGGCGGGCCATGCTCTGGGCACCTGGAG 1084
QY 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeu 360
DB 1085 TTAGATGACTTTTGGCCGGCTTCTCTCTGCAACCCAGGCGCATACCCCTCATCCAGAGCTA 1144
QY 361 ArgGlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLys 380
DB 1145 CGCAGGAACTGAGTCTTCCATACCTTGGCTTCCAGGACCCAGAGCTTGAAGTTCCAAA 1204
QY 381 ProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGln 400
DB 1205 CCAGGTGACGCTCTGAACCTGAGCATGGCCCGCCCGCTGGACAAGACACGTTCTGCCAG 1264
QY 401 GlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyrSerCysAla 420
DB 1265 GCGAAAGCTGATGGGCTCTATCCCAATCTCGGGAAACGCTCCAGCTTCTACAGCTGTGCA 1324
QY 421 AlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerAsnSerCysLys 440
DB 1325 GCGGGCGGCTGTTCCAGCAAGCTGCCCGACAGCGCTTGTGTTCAGCAACTCTCTGCAAA 1384
QY 441 CysCysThrTrpAsn 445
DB 1385 TGCTGCACCTGGAAT 1399

RESULT 3

US-09-267-574-1
; Sequence 1, Application US/09267574
; Patent No. 6399571
; GENERAL INFORMATION:
; APPLICANT: Gray, Patrick W.
; APPLICANT: Tjoelker, Larry W.
; TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
; FILE REFERENCE: 27866/35407
; CURRENT APPLICATION NUMBER: US/09/267,574
; CURRENT FILING DATE: 1999-03-12
; EARLIER APPLICATION NUMBER: 09/039,198
; EARLIER FILING DATE: 1998-03-12
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1636
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(1399)
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (65)..(1399)
US-09-267-574-1

Alignment Scores:

Pred. No.: 3 59e-238 Length: 1636
Score: 2398.00 Matches: 445
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-10-004-219B-10 (1-445) x US-09-267-574-1 (1-1636)

```
QY 1 AlaLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPhe 20
DB 65 GCAAACTGGTCTGCTACTTCAACAACTGGGCCAGTACAGACAGGGGAGGCTCGCTTC 124
QY 21 LeuProLysLeuAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet 40
DB 125 CTGCCAAGACTTGGACCCAGCCCTTTGACCACCTCATCTACGCCCTTCGCTGGCATG 184
QY 41 ThrAsnHisGlnLeuSerThrThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsn 60
DB 185 ACCAACCCAGCTGAGCACCACCTGAGTGGAGTACGAGACTCTCTACAGGAGTTCAAT 244
QY 61 GlyLeuLysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPhe 80
DB 245 GGCCCTGAAGAAGATGAATCCCAAGCTGAAGACCCTGTGTAGCCATCGAGGCTGGAATTC 304
QY 81 GlyThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsn 100
DB 305 GGCACTCAGAAAGTTACAGATATGGTAGCCACCGGCCAACACCGTCAGACCTTTGTCAAC 364
QY 101 SerAlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyr 120
DB 365 TCGCCCATCAGGTTCTTCGCCAATACAGCTTTGACGGCTTGACCTTGAGTGGAGTAC 424
QY 121 ProGlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAsp 140
DB 425 CCAGGAAGCCAGGGAGCCCTGCGTAGACAGAGCGCTTCAACACCTCTGGTACAGGAC 484
QY 141 LeuAlaAsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuSer 160
DB 485 TTGGCAATGCCCTTCAGCAGGAAGCCAGACCTCAGGAAGGAGCGCTTCTCTGAGT 544
QY 161 AlaAlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAla 180
DB 545 GCAGCGTTCCAGCTGGCAGACCTATGTGATGCTGTGATACAGGTGAGCAAAATCGCC 604
QY 181 GlnAsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysVal 200
DB 605 CAGAACCTGGATTTGTCAACCTTATGGCCTACGACTTCCATGGCTCTTGGGAGAGGTC 664
QY 201 ThrGlyHisAsnSerProLeuTyrLysArgGlnGluLysSerGlyValAlaAlaSerLeu 220
DB 665 ACGGAGCAATAACGCCCTCTACAGAGCAAGAGAGAGTGGTGAGCGCCCTC 724
QY 221 AsnValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuIle 240
DB 725 AACGTGATGCTGCTGTGCAACAGTGGCTGCAGAGGGAGCCCTGCCAGCAAGCTGATC 784
QY 241 LeuGlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgVal 260
DB 785 CTTGGCATGCTACCTACGAGCGCTCTTACACTGGCCCTCTCATCAGACACAGAGTG 844
QY 261 GlyAlaProAlaThrGlySerGlyThrProGlyProPheThrLysGlnGlyMetLeu 280
DB 845 GGGGCCCCAGCCAGGGTCTGGCATCTCAGGCCCTTACCAAGGAAGAGGAGTGCTG 904
QY 281 AlaTyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLys 300
DB 905 GCCTACTATGAGTCTGCTCTCGAAGGGGCCCAACACAGAGATCCAGGATCAGAAG 964
QY 301 ValProTyrIlePheArgAspAsnGlnTrpValGlyPheAspValGluSerPheLys 320
DB 965 GTGCCCTATCTTCGGGACAAACCAAGTGGGTGGGCTTTGATGATGTGAGAGCTTCAAA 1024
```

```
QY 321 ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyGlyAlaMetValTrpAlaLeuAsp 340
DB 1025 ACCAAGGTGAGTATCTGAGCAGAAAGGAGCTGGGGGGGCCCATGGTCTGGGCACTGGAC 1084
QY 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeu 360
DB 1085 TTAGATGACTTTTGGCGGCTTCTCTGCAACAGGCGCATACCCCTCATCCAGACGCTA 1144
QY 361 ArgGlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLys 380
DB 1145 CGCGCAGGAACCTGAGTCTTCCATACCTTCCAGCACCCAGAGCTTGAAGTTCCAAA 1204
QY 381 ProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGln 400
DB 1205 CCAGGTGACCCCTCTGACCTGAGCATGGCCCGCCCTGGACACAGACCTTCTGCCAG 1264
QY 401 GlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyrSerCysAla 420
DB 1265 GGCAGAGCTGATGGGCTCTATCCCAATCCTCGGAACGGTCCAGCTTCTACAGCTGTGCA 1324
QY 421 AlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerAsnSerCysLys 440
DB 1325 GCGGGGCGGCTGTTCCAGCAAGCTGCCCGACAGCGCTGTGTTCAGCAACTCCTGCAAA 1384
QY 441 CysCysThrTrpAsn 445
DB 1385 TGCTGCACCTGGAAT 1399
```

RESULT 4

US-08-486-839-3
; Sequence 3, Application US/08486839
; Patent No. 5928928
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: A human chitinase, its recombinant
; TITLE OF INVENTION: production, its use for decomposing chitin, its use
; TITLE OF INVENTION: in therapy or prophylaxis against infection diseases.
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11758
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,839
; FILING DATE: 07 - June - 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Baron, Ronald J.
; REGISTRATION NUMBER: 29,281
; REFERENCE/DOCKET NUMBER: 294-26
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1643 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; US-08-486-839-3

Alignment Scores:

Pred. No.: 9,376-238 Length: 1643
Score: 2394.00 Matches: 444
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.83% Indels: 0
DB: 2 Gaps: 0

US-10-004-219B-10 (1-445) x US-08-486-839-3 (1-1643)

```
QY 2 LysLeuValCysThrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPheLeu 21
DB 79 AAACGGGTCTGCTACTTACCAACCTGGGCCCGAGTACAGACAGGGGAGGCTCGCTTCCTG 138
QY 22 ProLysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMetThr 41
DB 139 CCCAAGGACTTGGACCCCGAGCCCTTGGACCCACCTCATCTAGCCCTTCGCTGGCATGACC 198
QY 42 AsnHisGlnLeuSerThrThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsnGly 61
DB 199 AACCAACGAGCTGAGCACCACCTGAGTGAATGACGAGACTCTCTACAGAGAGTTCATATGCC 258
QY 62 LeuLysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPheGly 81
DB 259 CTGAAGAAGATGAATCCCAAGCTGAAGACCCCTGTAGCCATCGGAGGCTTGGAAITTCGCG 318
QY 82 ThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsnSer 101
DB 319 ACTCAGAGATTACAGATATGTTAGCCAGCGCCACACACCTTGTCAACCTG 378
QY 102 AlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyrPro 121
DB 379 GCCATCAGGTTTCTCGCAAAATACAGCTTGTACGGCTTGACCTTGAGCTGGGAGTACCCA 438
QY 122 GlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAspLeu 141
DB 439 GGAAGCCAGGGAGGCGCTGCGCTAGCAAGAGGCGCTTCAACACCTGGTACAGACTG 498
QY 142 AlaAsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuSerAla 161
DB 499 GCCATGCTTCCAGAGGAGCCAGACCTCAGGAGGAGACGCTTCTTCTGAGTGCA 558
QY 162 AlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAlaGln 181
DB 559 GCGGTTCCAGCTGGGCGAGACCTATGTGGATGCTGATACAGAGTGGACAAAATGCCCGAG 618
QY 182 AsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysValThr 201
DB 619 AACCTGGATTGTCAACCTTATGGCTACGACTTCCATGGCTCTTGGGAGAGGTACG 678
QY 202 GlyHisAsnSerProLeuTyrLysArgGlnGluSerGlyAlaAlaSerLeuAsn 221
DB 679 GGACATACAGCCCTCTACAGAGGCGAGAGAGAGTGGTGCAGACCCAGCTCAAC 738
QY 222 ValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuLeu 241
DB 739 GTGGATGCTGCTGTGCAACAGTGGCTGCAGAGGGGACCCCTGCCAGCAAGCTGATCCTT 798
QY 242 GlyMetProThrTyrGlyArgSerPheThrIleuAlaSerSerAspThrArgValGly 261
DB 799 GGCATGGCTTACCTACGAGCGCTTCTTCACTGGCTTCTCATCAGACACCCAGAGTGGGG 858
QY 262 AlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyGlyMetLeuAla 281
DB 859 GCCCAGCCAGCAGGCTTGGCACTCAGGCCCTTCAACAGAGAGAGGAGGATGCTGGCC 918
QY 282 TyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLysVal 301
DB 919 TACTATGAATCTGCTCTCTGGAAGGGGGCCACCAACAGAGAGATCCAGGATCAGAAGGTG 978
QY 302 ProTyrIlePheArgAspAsnGlnTrpValGlyPheAspValGluSerPheLysThr 321
DB 979 CCCTACATCTCCGGGACAAACAGTGGGTGGGCTTGTGATGTGGAGAGCTTCAAAACC 1038
```

```
QY 322 LysValSerTyrLeuLysGlnLysGlyLeuGlyGlyAlaMetValTrpAlaLeuAspLeu 341
DB 1039 AAGGTCAGCTATCTGAAGCAGAGGAGCTGGGGGGCCATGGTCTGGGCACTGGACTTA 1098
QY 342 AspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeuArg 361
DB 1099 GATGACTTTGGCGGCTTCTCTCTCAACAGGAGGCGGATACCCCTCATCCAGACGCTACGG 1158
QY 362 GlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLysPro 381
DB 1159 CAGGACAGTGTCTTCATCTTGCCTTCAGGACCCCGAGCTTGAAGTTCCAAACCA 1218
QY 382 GlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGlnGly 401
DB 1219 GGTCAAGCCCTCTGAACCTGAGCAGTGGCCCGCCAGCCCTGGACAAGACACGCTTCTGCCAGGGC 1278
QY 402 LysAlaAspGlyLeuTyrProAsnProArgGlnArgSerSerPheTyrSerCysAlaAla 421
DB 1279 AAAGCTGATGGGCTCTATCCCAATCTCTCGGACGCTCCAGCTTCTACAGCTGTGCAGCG 1338
QY 422 GlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerAsnSerCysLysCys 441
DB 1339 GGGCGGCTGTTCCAGCAAAAGCTGCCGACAGGCTGGTGTTCAGCAACTCTCTGCAAAATGC 1398
QY 442 CysThrTrpAsn 445
DB 1399 TGCACCTGGAAT 1410
```

RESULT 5

US-09-151-011-3
; Sequence 3, Application US/09151011
; Patent No. 6057142

GENERAL INFORMATION:

APPLICANT:
TITLE OF INVENTION: A Human Chitinase, Its Recombinant
TITLE OF INVENTION: Production, Its Use For Decomposing Chitin, Its Use in
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hoffmann & Baron, LLP
STREET: 6900 Jericho Turnpike
CITY: Syosset
STATE: New York
COUNTRY: United States of America
ZIP: 11791

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/151,011
FILING DATE: 10 - September - 1998
ATTORNEY/AGENT INFORMATION:
NAME: Morris, Robert C.
REGISTRATION NUMBER: 42,910
REFERENCE/DOCKET NUMBER: 294-32 DIV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 822-3550
TELEFAX: (516) 822-3582
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1643 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO

US-09-151-011-3

Alignment Scores:

Pred. No.: 9,376-238 Length: 1643

Score: 2394.00 Matches: 444
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.83% Indels: 0
DB: 3 Gaps: 0

US-10-004-219B-10 (1-445) x US-09-151-011-3 (1-1643)

QY 2 LysLeuValCysThrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPheLeu 21
DB 79 AACTGGTCTGCTACTTCAACAACTGGGCCAGTACAGACAGGGGAGGCTGCTTCCTG 138
QY 22 ProLysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMetThr 41
DB 139 CCCAAGGACTTGGACCCCGAGCCTTTGCACCCACCTCATCTAGCCCTTGCCTGGCATGCC 198
QY 42 AsnHisGlnLeuSerThrThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsnGly 61
DB 199 AACCCACGAGCTGAGCACCACTGAGTGGAAATGACGAGACTCTCTACCGAGGATTCAATGGC 258
QY 62 LeuLysLeuMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPheGly 81
DB 259 CTGAAGAAGATGAATCCCAAGCTGAAGACCCCTTTAGCCATCGAGGCTGGAAATTTGGGC 318
QY 82 ThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsnSer 101
DB 319 ACTCAGAAGTTACACATATGGTAGCCACCGCCCAACACCGTCAGACCTTTGTCAACTCG 378
QY 102 AlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyrPro 121
DB 379 GCCATCAGGTTTCTGGCGAAATACACGCTTGGAGCGCTTGACCTTGACCTGGGAGTACCCA 438
QY 122 GlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAspLeu 141
DB 439 GGAAGCCAGGGAGCCCTCCGCTAGACAGAGCGCTTCAACACCTGGTACAGGACTTG 498
QY 142 AlaAsnAlaPheGlnGlnGlnAlaGlnThrSerGlyLysGluArgLeuLeuSerAla 161
DB 499 GCCAATGCCTTCCAGCAGGAAGCCAGACCTCAGGGAAGAACGCTTCTCTGAGTGCA 558
QY 162 AlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAlaGln 181
DB 559 GCGGTTCCAGCTGGGAGACCTATGTGGATGCTGGATACAGGTGGACAAATCGCCAG 618
QY 182 AsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysValThr 201
DB 619 AACCTGGATTTGTCAACCTTATGGCTTACGACTTCATGGCTCTTGGAGAGGTCACG 678
QY 202 GlyHisAsnSerProLeuTyrLysArgGlnGluSerGlyAlaAlaSerLeuAsn 221
DB 679 GGACATAACAGCCCTCTACAGAGGCAAGAGAGAGTGGTGCAGCAGCCCTCAAC 738
QY 222 ValAspAlaAlaValGlnGlnThrLeuGlnLysGlyThrProAlaSerLysLeuIleLeu 241
DB 739 GTGGATGCTGCTGTGCAACAGTGGCTGCAGAGGGGACCCCTGCCAGCAAGCTGATCCTT 798
QY 242 GlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgValGly 261
DB 799 GGCATGCCCTACCTACGAGCGCTCCITTCACACTGGCTCTCATCAGACACAGAGTGGGG 858
QY 262 AlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyMetLeuAla 281
DB 859 GCCCCAGCCACAGGCTTGGCACTCCAGGCCCTTCAACAGGAAGAGGAGGATGCTGGCC 918
QY 282 TyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLysVal 301
DB 919 TACTATGAGTCTGCTCCTGGAGGGGGGCCACCAACAGAGATCCAGGATCAGAAGGTG 978
QY 302 ProTyrIlePheArgAspAsnGlnTrpValGlyPheAspValGluSerPheLysThr 321
DB 979 CCTCATCTTCCGGGACAAACAGTGGTGGCTTTGATGATGGAGAGCTTCAAAACC 1038
QY 322 LysValSerTyrLeuLysGlnLysGlyLeuGlyAlaMetValTrpAlaLeuAspLeu 341

DB 1039 AAGGTCAAGTATCTCAAGCAGAAAGGAGCTGGGGCGGGCCATGGTCTGGGCACTTGA 1098
QY 342 AspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeuArg 361
DB 1099 GATGACTTTGGCGGCTTCTCTGCAACAGGCGCCATACCCCTCATCCAGAGCTACGG 1158
QY 362 GlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLysPro 381
DB 1159 CAGGAACCTGAGTCTTCCATACCTTGCCTTCAGGACCCCGAGAGCTTGAAGTTC 1218
QY 382 GlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGlnGly 401
DB 1219 GGTCAAGCCTCTGAACCTGAGCATGGCCCGCCAGCCCTGGACACACGCTTCTGCCAGG 1278
QY 402 LysAlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyrSerCysAlaAla 421
DB 1279 AAAGCTGATGGGCTCTATCCCAATCTCGGAAACGCTCCAGCTTCTACAGCTGTG 1338
QY 422 GlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerAsnSerCysLysCys 441
DB 1339 GGGCGGCTTCTCCAGCAAGGCTGCCCGACAGGCTGGTGTTCAGCAACTCTCTGCAATGC 1398
QY 442 CysThrTrpAsn 445
DB 1399 TGCACCTGGAAT 1410

RESULT 6

US-09-343-623-3
; Sequence 3, Application US/09343623
; Patent No. 6303118

GENERAL INFORMATION:

APPLICANT: A human chitinase, its recombinant
TITLE OF INVENTION: production, its use for decomposing chitin, its use
TITLE OF INVENTION: in therapy or prophylaxis against infection diseases.
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:

ADDRESSER: Hoffmann & Baron
STREET: 350 Jericho Turnpike
CITY: Jericho

STATE: New York

COUNTRY: United States of America

ZIP: 11758

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

SOFTWARE:

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/343,623

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/486,839

FILING DATE: 07-June-1995

ATTORNEY/AGENT INFORMATION:

NAME: Baron, Ronald J.

REGISTRATION NUMBER: 29,281

REFERENCE/DOCKET NUMBER: 294-26

TELECOMMUNICATION INFORMATION:

TELEPHONE: (516) 822-3550

TELEFAX: (516) 822-3582

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 1643 base pairs

TYPE: nucleic acid

STRANDEDNESS: unknown

TOPOLOGY: unknown

MOLECULE TYPE: cDNA

HYPOTHETICAL: NO

US-09-343-623-3

Alignment Scores:

Pred. No.: 9,37e-238 Length: 1643
Score: 2394.00 Matches: 444
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.83% Indels: 0
DB: 4 Gaps: 0

US-10-004-219B-10 (1-445) x US-09-343-623-3 (1-1643)

```
QY 2 LysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPheLeu 21
DB 79 AAACGGGTCTGCTACTTACCAACATGGGCCCACTAGTACAGACAGGGGAGGCTGCTTCCTG 138
QY 22 ProLysAspLeuAspProSerLeuCysThrHisLeuLeuLeuTyrAlaPheAlaGlyMetThr 41
DB 139 CCCAAGGACTTGGACCCAGCCCTTGGACCCACCTCCTACGCTTCTGCTGGCATGACC 198
QY 42 AsnHisGlnLeuSerThrThrGluTrpAsnAspGluThrLeuTyrGlnLeuPheAsnGly 61
DB 199 AACCAACAGCTGAGCACCACCTGAGTGGATGACGAGACTCTCTACACGAGTTCAATGGC 258
QY 62 LeuLysLysMetAsnProLysLeuLysThrLeuLeuAlaLeuGlyClyTrpAsnPheGly 81
DB 259 CTGAAGAAGATGAATCCCAAGCTGAAGACCTCTTGTAGCCATCGGAGGCTGGAAATTCGCG 318
QY 82 ThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsnSer 101
DB 319 ACTCAAGATTCACAGATATGTTAGCCAGGCGCAACACCTGACAGCTTTGTCAACTCG 378
QY 102 AlaLeuArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyrPro 121
DB 379 GCCATCAGGTTCTCGCCAAATACAGCTTTGACGCGCTTACCTTGTGAGTACCCA 438
QY 122 GlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAspLeu 141
DB 439 GGAAGCCAGGGAGGCGCTCGCCGTAGCAAGGAGCGCTTCACAAACCTGCTGACAGACTG 498
QY 142 AlaAsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuSerAla 161
DB 499 GCCAATGCTTCCAGCAGGAGCCAGACCTCAGGAGAGAGAGAGCTTCTTCTGAGTGCA 558
QY 162 AlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysLeuAlaGln 181
DB 559 GCGGTTCCAGCTGGGACAGACCTATGTGATGCTGATGAGTGCGACAAAATGCCAC 618
QY 182 AsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysValThr 201
DB 619 AACCTGGATTTTGTCAACCTTATGSCCTACGACTTCCATGGCTCTTGGGAGAAGTCCAG 678
QY 202 GlyHisAsnSerProLeuTyrLysArgGlnGluGluSerGlyValAlaAlaSerLeuAsn 221
DB 679 GGACATAACAGCCCTCTACAAAGGCGAAGAGAGAGTGGTGACAGCCAGCCCTCAAC 738
QY 222 ValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuLeuLeu 241
DB 739 GTGGATGCTGCTGCAACAGTGGCTGCAAGAGGAGCCCTGCCAGCAAGCTGATCCTT 798
QY 242 GlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgValGly 261
DB 799 GGCATGGCTTACTACGAGCGCTCTCTTCACTGGCTCTCTATCAGACACCAAGAGTGGG 858
QY 262 AlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyMetLeuAla 281
DB 859 GCCCCAGGCACAGGGTCTGGCACTCCAGGCCCTTCCAAAGAAAGAGGAGGATGCTGGCC 918
QY 282 TyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLysVal 301
DB 919 TACTATGAAGTCTGCTCTCTGGAAGGGGCCCAACCAACAGAGATCCAGGATCAGAGGTG 978
QY 302 ProTyrIlePheArgAspAsnGlnTrpValGlyPheAspValGluSerPheLysThr 321
DB 979 CCCTACATCTTCCGGGACAACCAAGTGGGTGGGCTTTGATGATGGAGAGCTTCAAAACC 1038
```

```
QY 322 LysValserTyrLeuLysGlnLysGlyLeuGlyGlyAlaMetValTrpAlaLeuAspLeu 341
DB 1039 AAGGTACGTATCTGAGACAGAGGACTGGCGGGGCGCATGGTCTGGGCACCTGACTTA 1098
QY 342 AspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeuArg 361
DB 1099 GATGACTTTGGCGGCTTCTCTGCAACCCAGGGCGGATACCCCTCATCCAGACGCTACGG 1158
QY 362 GlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLysPro 381
DB 1159 CAGGAACGTAGTCTTCCATCTTGCCTTTCAGCACCCCGAGAGCTTGAAGTTCCAAAACA 1218
QY 382 GlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGlnGly 401
DB 1219 GGTCAAGCCCTCTGAACCTGAGCATGGCCCGAGCCCTGACCAAGACACCTTCTGCCAGGGC 1278
QY 402 LysAlaAspGlyLeuTyrProAsnProArgGluArgSerPheTyrSerCysAlaAla 421
DB 1279 AAGCTGATGGCTCTATCCCAATCTCGGGAACGGTCCAGCTTCTACAGCTGTGCAGCG 1338
QY 422 GlyArgLeuPheGlnGlnSerCysProThrThrGlyLeuValPheSerAsnSerCysLysCys 441
DB 1339 GGGCGGCTGTTCAGCAAAAGCTGCCCGACAGGCCCTGGTGTTCAGCAACTCTCTGCAATGC 1398
QY 442 CysThrTrpAsn 445
DB 1399 TGCACCTGGAAAT 1410

RESULT 7
US-09-039-198A-3
; Sequence 3, Application US/09039198A
; Patent No. 6200951
; GENERAL INFORMATION:
; APPLICANT: Gray, Patrick W.
; APPLICANT: Tjoelker, Larry W.
; TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive/6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/039,198A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 27866/34391
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1656 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 27..1424
; FEATURE:
; NAME/KEY: mat_peptide
```

LOCATION: 90..1424

US-09-039-198A-3

Alignment Scores:

Score: 1.53e-237 Length: 1656
Matches: 444
Conservative: 0
Percent Similarity: 99.78%
Best Local Similarity: 99.78%
Query Match: 99.75%
Indels: 0
Gaps: 0
DB: 4

US-10-004-219b-10 (1-445) x US-09-039-198A-3 (1-1656)

QY 1 AlaLysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPhe 20
DB 90 GCAAAACTGGTCTGCTACTTCCAACTGGGCCAGTACAGACAGGGGAGGCTCGCTTC 149
QY 21 LeuProLysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet 40
DB 150 CTGCCCAAGGACTTGGACCCAGCCCTTTGCACCCACTCATCTACGCCCTTCGCTGGCATG 209
QY 41 ThrAsnHisGlnLeuSerThrThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsn 60
DB 210 ACCAACCCAGCAGCTGAGCACCCTGAGTGGATGGAATGACGAGACTCTTACCCAGGAGTTCAAT 269
QY 61 GlyLeuLysLysMetAsnProLysLeuLysThrLeuAlaIleGlyGlyTrpAsnPhe 80
DB 270 GGCTGAGAGAGATGATCCCAAGCTGAGACCCCTGTAGCCATCGAGGCTGGAAATTC 329
QY 81 GlyThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsn 100
DB 330 AGCACTCAGAAGTTACAGATATGTTAGCCAGCCCAACCAACCGTCAGACCTTTGTCAAC 389
QY 101 SerAlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyr 120
DB 390 TCGGCCATCAGGTTCTTGGCAATACAGCTTTGACGGCTTGACCTTACCTGGAGTAC 449
QY 121 ProGlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAsp 140
DB 450 CCAGGAGCCAGGGGAGCCCTGCGGTAGACAGGAGCGCTTACACACCTCTGTACAGGAC 509
QY 141 LeuAlaAsnAlaPheGlnGlnGlnAlaGlnThrSerGlyLysGluArgLeuLeuSer 160
DB 510 TTGGCCATGCTTCCAGCAGGAGCCAGACCTCAGGGAAGAACGCTTCTTCTGAGT 569
QY 161 AlaAlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAla 180
DB 570 GCAGCGGTTCCAGCTGGGAGACCTATGTTGGATGCTGGATACAGGTGGACAAAAATCGCC 629
QY 181 GlnAsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysVal 200
DB 630 CAGNACTGGATTTGTCAACCTTATGGCTAGCACTTCCATGGCTCTTGGGAGAGGTC 689
QY 201 ThrGlyHisAsnSerProLeuTyrLysArgGlnGlnGluSerGlyAlaAlaSerLeu 220
DB 690 ACGGACATAACAGCCCTCTACAAGAGCAGAGAGAGAGTGTGTCAGCAGCCAGCTTC 749
QY 221 AsnValAspAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuIle 240
DB 750 AACGTGGATGCTGCTGCAACAGTGGCTGCAAGAGGGGACCCCTGCCAGCAAGCTGATC 809
QY 241 LeuGlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgVal 260
DB 810 CTGGGATGCTTACCTACGAGCGCTCTTCACTGAGCTCTTCACTACAGCAGGATG 869
QY 261 GlyAlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyMetLeu 280
DB 870 GGGGCCCCAGCCAGGGCTGCGCACTCCAGGCCCCCTTCAACAAAGAGAGGAGGATGCTG 929
QY 281 AlaTyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLys 300
DB 930 GCCTACTATGAAGTCTGCTCTTGGAGGGGGGCCCAACCAACAGAGATCCAGGATCAGAAG 989

QY 301 ValProTyrIlePheArgAspAsnGlnTrpValGlyPheAspAspValGluSerPheLys 320
DB 990 GTGCCCTACATCTTCCGGGCAACACCACTGGGTGGGCTTTGATGATGTGGAGAGCTTCAAA 1049
QY 321 ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyGlyAlaMetValTrpAlaLeuAsp 340
DB 1050 ACCAAGGTCAGCTATCTGAAGCAGAAGGACTGGGGGGGGCCCATGGTCTGGGCACTGGAC 1109
QY 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeu 360
DB 1110 TTAGATGACTTTCGGGCTTCTCTGCAACCAAGGCCCATACCCCTCATCCAGAGCTA 1169
QY 361 ArgGlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLys 380
DB 1170 CGGCAGGAACCTGAGTCTTCCATCTTGGCTTCCAGGACCCCAAGAGCTTGAAGTTCACAAA 1229
QY 381 ProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGln 400
DB 1230 CCAGGTCAAGCCCTCTGAACCTGAGCATGGCCCAAGCCCTGGCAAGACACGTTCTGCCAG 1289
QY 401 GlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyrSerCysAla 420
DB 1290 GCAAAAGCTGATGGCTCTATCCCAATCTCGGGAACGTCAGCTTCTACAGCTGTGCA 1349
QY 421 AlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerAsnSerCysLys 440
DB 1350 GCGGGGGCGCTTCTCAGCAAGCTGCCCAAGGCTGCCCAAGGCTGGTGTTCAGCAACTCTCTGCAAA 1409
QY 441 CysCysThrTrpAsn 445
DB 1410 TGCTGCACCTGGAAT 1424

RESULT 8
US-08-877-599-3
; Sequence 3, Application US/08877599
; Patent No. 6372212
; GENERAL INFORMATION:
; APPLICANT: Gray, Patrick W.
; TITLE OF INVENTION: Chitinase Materials and Methods
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/877,599
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/663,618
; FILING DATE: 14-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 27866/33994
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1656 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 27..1424
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 90..1424
US-08-877-599-3

Alignment Scores:

Pred. No.: 153e-237 Length: 1656
Score: 2392.00 Matches: 444
Percent Similarity: 99.78% Conservative: 0
Best Local Similarity: 99.78% Mismatches: 1
Query Match: 99.75% Indels: 0
DB: 4 Gaps: 0

US-10-004-219B-10 (1-445) x US-08-877-599-3 (1-1656)

QY 1 AlaLysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyAlaAlaArgPhe 20
DB 90 GCAAAACTGGTCTGCTACTTCCAACTGGGCCAGTACAGACAGGGGAGGCTGGCTTC 149
QY 21 LeuProLysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet 40
DB 150 CTGCCCAAGGACTTGGACCCAGCCTTTGCACCACCTCATCTACGCTTTCGCTGGCATG 209
QY 41 ThrAsnHisGlnLeuSerThrThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsn 60
DB 210 ACCAACCCAGCTGAGCAGCCTGAGTGGTGAATGACGAGACTCTTACCAGGAGTTCAAT 269
QY 61 GlyLeuLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPhe 80
DB 270 GGCCTGAAGAGATGAATCCCAAGCTGAAGACCTCTTAGCCCATCGGAGGCTGGAATTC 329
QY 81 GlyThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsn 100
DB 330 AGCACTCAGAAGTTCCAGATATGGTAGCCAGCCGCAACCCGTCAGACCTTTGTCAAC 389
QY 101 SerAlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyr 120
DB 390 TGGCCATCAGGTTCTTGGCAATACAGCTTTGACGGCTTGACCTTGACCTGGGAGTAC 449
QY 121 ProGlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAsp 140
DB 450 CCAGGAAGCCAGGGAGCCCTGCGTAGACAGAGGCTTCAACCTCTGGTACAGGAC 509
QY 141 LeuAlaAsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuSer 160
DB 510 TTGGCCAATGCCCTTCCAGCAGGAAGCCAGACCTCAGGGAAGGAGACGCTTCTTCTGAGT 569
QY 161 AlaAlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAla 180
DB 570 GCAGCGGTTCAGCTGGGAGAGCTATGGATGGTGTACAGGTGACAAAATCGCC 629
QY 181 GlnAsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysVal 200
DB 630 CAGAACTGGATTGTCAACCTTATGGCTACAGCTTCCATGGCTCTTGGGAGAGGTC 689
QY 201 ThrGlyHisAsnSerProLeuTyrLysArgGlnGluSerGlyAlaAlaAlaSerLeu 220
DB 690 ACGGACATAACAGCCCTCTTCAAGAGGCAAGAGAGAGTGGTGCAGCAGCCAGCTC 749
QY 221 AsnValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuIle 240
DB 750 AAGTGGATGCTGCTGTGCAACAGTGGCTGCAGAGGGGACCCCTGCCAGAGCTGATC 809
QY 241 LeuGlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgVal 260
DB 810 CTTGGCATGCTACTACGAGCGCTCTTCCACTGGCTCTCATCAGACACAGAGTG 869
QY 261 GlyAlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyMetLeu 280

DB 870 GGGGCCCCAGCCACAGGCTCTGGCACTCCAGGCCCTTCCAAAGAAAGGAGGATGCTG 929
QY 281 AlaTyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLys 300
DB 930 GCCTACTATGAAGTCTGCTCTGGNAGGGGGCCCAACACAGAGATCCAGGATCAGAAG 989
QY 301 ValProTyrIlePheArgAspAsnGlnTrpValGlyPheAspValGluSerPheLys 320
DB 990 GTGCCCTACATCTTCCGGGCAACCCAGTGGGTGGGCTTTGATGATGTGGAGAGCTTCAAA 1049
QY 321 ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyGlyAlaMetValTrpAlaLeuAsp 340
DB 1050 ACCAAGGTCAGCTATCTGAAGCAGAGGAGCTGGGCGGGCCCATGGTCTGGGCACCTGGAC 1109
QY 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeu 360
DB 1110 TTAGATGACTTTGCGGCTTCTCTGCAACCCAGGCGCATACCCCTCATCCAGAGCTA 1169
QY 361 ArgGlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLys 380
DB 1170 CGGCAGGAACCTGAGTCTTCCATCTTGGCTTTCAGGACCCAGAGCTTGAAGTTCCAAA 1229
QY 381 ProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGln 400
DB 1230 CCAGGTGACGCCCTCTGAACCTGAGCATGGCCCGCCCTGGACAAGACACGTTCTGCCAG 1289
QY 401 GlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyrSerCysAla 420
DB 1290 GGCAGAGCTGATGGCTCTATCCCAATCTCTGGGAACGGTCCAGCTTCTACAGCTGTGCA 1349
QY 421 AlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerAsnSerCysLys 440
DB 1350 GCGGGCGGCTGTTCAGCAAGAGCTGCCGACAGGCTGGTGTTCAGCAACTCTCTGCAAA 1409
QY 441 CysCysThrTrpAsn 445
DB 1410 TGCTGCACCTGGAAT 1424

RESULT 9

US-09-267-574-3
Sequence 3, Application US/09267574
Patent No. 639571
GENERAL INFORMATION:
APPLICANT: Gray, Patrick W.
APPLICANT: Tjoelker, Larry W.
TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
FILE REFERENCE: 27866/35407
CURRENT APPLICATION NUMBER: US/09/267,574
CURRENT FILING DATE: 1999-03-12
EARLIER APPLICATION NUMBER: 09/039,198
EARLIER FILING DATE: 1998-03-12
NUMBER OF SEQ ID NOS: 39
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
LENGTH: 1656
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (27)..(1424)
FEATURE:
NAME/KEY: mat_peptide
LOCATION: (90)..(1424)
US-09-267-574-3

Alignment Scores:
Pred. No.: 1.53e-237 Length: 1656
Score: 2392.00 Matches: 444
Percent Similarity: 99.78% Conservative: 0
Best Local Similarity: 99.78% Mismatches: 1
Query Match: 99.75% Indels: 0
DB: 4 Gaps: 0

US-10-004-219B-10 (1-445) x US-09-267-574-3 (1-1656)

QY 1 AlaLysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPhe 20
 Db 90 GCAAACTGGTGTCTACTTCAACCACTGGGCCAGTACAGACAGGGGGAGGCTCGCTTC 149
 QY 21 LeuProLysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet 40
 Db 150 CTGCCCAAGGACTTGGACCCAGCTTTGGACCCACCTCATCTACGGCTTCGGTGGCATG 209
 QY 41 ThrAsnHisGlnLeuSerThrThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsn 60
 Db 210 ACCAACCCAGCTGAGCACCCTGAGTGGGAATCAGCAGACTCTTACACGAGGATCAAT 269
 QY 61 GlyLeuLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTyrAsnPhe 80
 Db 270 GGCTTGAAGAAGATGAATCCAGCTGAAGACCTGTGTAGCCATCGGAGGCTGGAATTC 329
 QY 81 GlyThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsn 100
 Db 330 AGCACTCAGAAGTTTCAGATATGGTAGCCACCGCCCAACACCGTCAGACTTTGTCAAC 389
 QY 101 SerAlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyr 120
 Db 390 TCGGCCATCAGGTTTCTGCGCAATACAGCTTTGACGGCCTTGACCTTGACTGGGAGTAC 449
 QY 121 ProGlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAsp 140
 Db 450 CCAGGAAGCAGGAGGAGCCCTGCGGTAGACAGAGCGCTTCAACACCTGGTACAGGAC 509
 QY 141 LeuAlaAsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuLeuSer 160
 Db 510 TTGGCCATCCCTTCCAGCAGGAGGCCACACCTCAGGAGGAGGACCGCTTCTTGAGT 569
 QY 161 AlaAlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAla 180
 Db 570 GCAGCGGTTCAGCTGGCCAGACCTATGTGGATGTGGATACGAGGTGGACAAATCGCC 629
 QY 181 GlnAsnLeuAspPheValAsnLeuMetAlaTyrAsnPheHisGlySerTrpGluLysVal 200
 Db 630 CAGAACCTGGATTTTGAACCTTATGGCTAGACTTCCATGGCTCTTGGGAGAGGTC 689
 QY 201 ThrGlyHisAsnSerProLeuTyrLysArgGlnGluSerGlyAlaAlaLaserLeu 220
 Db 690 ACGGACATTAACGCCCTCTACAGAGGCAAGAGAGTGTGGCAGCAGCAGCCCTC 749
 QY 221 AsnValAspAlaValGlnGlnTrpLeuGlnLysGlyThrProLaserLysLeuIle 240
 Db 750 AACGTGGATGCTGTGTGCAACAGTGGCTGCAGAAAGGGGACCCCTGCGCAGCAAGCTGATC 809
 QY 241 LeuGlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgVal 260
 Db 810 CTGGGATGCTTACCTACGAGCGCTCTTCACTAGTGGCTTCCATCAGACACAGAGTG 869
 QY 261 GlyValaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyGlyMetLeu 280
 Db 870 GGCGCCCAACCCAGGGTCTGGCACTCCAGGCCCTTCAACAGGAGGAGGAGTGGCTG 929
 QY 281 AlaTyrTrpGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLys 300
 Db 930 GCCTACTATGAAGTCTGCTCTGGAAGGGGGCCACCAACAGAGAATCCAGGATCAAG 989
 QY 301 ValProTyrIlePheArgAspAsnGlnTrpValGlyPheAspValGluSerPheLys 320
 Db 990 GTGCCCTTACATCTTCCGGGCAACACCTGGTGGCTTTGATGTGGAGAGCTTCAAA 1049
 QY 321 ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyAlaMetValTrpAlaLeuAsp 340
 Db 1050 ACCAAGTCACTATCTGAAGCAGAGGAGTGGCGGGCCATGTCTGGGCACTGGAC 1109
 QY 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeu 360
 Db 1110 TTAGATGACTTTTCCGGCTTCTCTGTCAACACGAGGGCGGATACCCCTCTATCCAGAGCTA 1169

QY 361 ArgGlnGluLeuSerLeuProTyrTrpLeuProSerGlyThrProGluLeuGluValProLys 380
 Db 1170 CGCAGGAACTGAGTCTTCCATCTTGCCTTCAGGACCCAGAGCTTGAAGTTCCAAA 1229
 QY 381 ProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGln 400
 Db 1230 CCAGGTGACCCCTCTGAACCTGAGCATGSCCCAGCCCTGGACAAGACACGCTTCTGCCAG 1289
 QY 401 GlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyrSerCysAla 420
 Db 1290 GGCAGGCTGATGGGCTCTATCCCAATCCTCGGAAACGGTCCAGCTTCTACAGCTGTGCA 1349
 QY 421 AlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerAsnSerCysLys 440
 Db 1350 GCGGGGCGGCTGTTCAGCAAGAGCTGCCGACAGGCTGGTGTTCAGCAACTCTCTGCAAA 1409
 QY 441 CysCysThrTrpAsn 445
 Db 1410 TGTGCACTGGAAT 1424

RESULT 10
 US-08-486-839-5
 ; Sequence 5, Application US/08486839
 ; Patent No. 5928928
 ; GENERAL INFORMATION:
 ; APPLICANT:
 ; TITLE OF INVENTION: A human chitinase, its recombinant
 ; TITLE OF INVENTION: production, its use for decomposing chitin, its use
 ; TITLE OF INVENTION: in therapy or prophylaxis against infection diseases
 ; NUMBER OF SEQUENCES: 16
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Hoffmann & Baron
 ; STREET: 350 Jericho Turnpike
 ; CITY: Jericho
 ; STATE: New York
 ; COUNTRY: United States of America
 ; ZIP: 11758
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/486,839
 ; FILING DATE: 07 - June - 1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Baron, Ronald J.
 ; REGISTRATION NUMBER: 29,281
 ; REFERENCE/DOCKET NUMBER: 294-26
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (516) 822-3550
 ; TELEFAX: (516) 822-3582
 ; INFORMATION FOR SEQ ID NO: 5:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1713 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: unknown
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: cDNA
 ; HYPOTHETICAL: NO
 ; US-08-486-839-5

Alignment Scores:
 Pred. No.: 5,048-233 Length: 1713
 Score: 2348.50 Matches: 443
 Percent Similarity: 94.66% Conservative: 0
 Best Local Similarity: 94.66% Mismatches: 1
 Query Match: 97.94% Indels: 25
 DB: 2 Gaps: 1
 US-10-004-219B-10 (1-445) x US-08-486-839-5 (1-1713)

QY 2 LysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPheLeu 21
DB 79 AAACGGTCTGCTACTTACCAACTGGGCCAGTACAGACAGGGGAGGCTCGCTTCCTG 138
QY 22 ProLysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMetThr 41
DB 139 CCCAAGGACTTGGACCCCAAGCCCTTTGCACCCACCTCATCTAGCCCTTCGCTGGCATGACC 198
QY 42 AsnHisGlnLeuSerThrThrGlnTrpAsnAspGluThrLeuTyrGlnGluPheAsnGly 61
DB 139 AACCAACAGCTAGACACCACTGAGTGGATGAGAGACTCTCTACAGAGGTCAATGGC 258
QY 62 LeuLysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyTyrAsnPheGly 81
DB 259 CTGAAGAAGATGAATCCCAAGCTGAAGACCCCTGTAGCCATCGGAGGCTGGAATTCGGC 318
QY 82 ThrGlnLysPheThrAspMetValIleAlaAsnAsnArgGlnThrPheValAsnSer 101
DB 319 ACTCAGAAGTTTCAACAGATATGTTAGCCACCGGCAACAACCGTCAGACCTTTGTCAACTCG 378
QY 102 AlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGlnTyrPro 121
DB 379 GCCATCAGTTTCTGCGCAATACAGCTTTGACGGCTTTGACCTTGACTGGGAGTACCCA 438
QY 122 GlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAspLeu 141
DB 439 GGAAGCCAGGGAGGCCCTGCGCTAGCAAGAGAGCGCTTCAACACCTGGTACAGACTTG 498
QY 142 AlaAsnAlaPheGlnGlnGlnAlaGlnThrSerGlyLysGluArgLeuLeuSerAla 161
DB 499 GCCAATGCTTTCAGCAGGAAGCCAGACCTCAGGGAAGAAAGCGCTTCTTCTGAGTGCA 558
QY 162 AlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAlaGln 181
DB 559 GGGGTTCAGCTGGGAGACCTATGTGGATGCTGGATAGAGTGGACAAAATGCCACG 618
QY 182 AsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysValThr 201
DB 619 AACCTGGATTTTGTCAACCTTATGCGCTACGACTCCATGGCTCTTGGGAGAGGTCAAG 678
QY 202 GlyHisAsnSerProLeuTyrLysArgGlnGlnGluSerGlyAlaAlaSerLeuAsn 221
DB 679 GGACATAACAGCCCTCTTACAAGAGGCAAGAGAGAGTGGTGACAGCCAGCCAGCTCAAC 738
QY 222 ValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuIleLeu 241
DB 739 GTGGATGCTGCTGTGCAACAGTGGCTGCAAGAGGGGACCCCTGCCAGCAAGCTGATCCTT 798
QY 242 GlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerAspThrArgValGly 261
DB 799 GGCATGCTTACCTACGGAGGCTCTCTCACACTGGCTCTCTCATCAGACACAGAGTGGG 858
QY 262 AlaProAlaThrGlySerGlyThrProGlyProPheThrLysGlnGlyGlyMetLeuAla 281
DB 859 GCGCCAGCCACAGGGTCTGGCACTCAGGCCCTTCCACCAAGAGAGGGATGCTGGCC 918
QY 282 TyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLysVal 301
DB 919 TACTATGAAGTCTGCTCTGGAAAGGGGGCCACCAACAGAGAATCCAGGATCAGAGGTG 978
QY 302 ProTyrIlePheArgAspAsnGlnTrpValGlyPheAspValGluSerPheLysThr 321
DB 979 CCTCATCTCTCCGGACACCAACAGTGGGTGGCTTTGATGATGGAGAGCTTCAAAACC 1038
QY 322 LysValSerTyrLeuLysGlnLysGlyLeuGlyAlaMetValTrpAlaLeuAspLeu 341
DB 1039 AAGGTCAAGTATCTGAAGCAGAAAGGAGTGGGGGGCCCATGCTGGGCACTGGACTTA 1098
QY 342 AspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeuArg 361
DB 1099 GATGACTTTGGCCGGCTTCTCTGCAACCAAGGGCCGATACCCCTCATCCAGACGCTACGG 1158

QY 362 GlnGluLeuSer----- 365
DB 1159 CAGGAACCTGAA-TGGGTAAAGCCTTAACCTGCTGTACATGTGAGGCCAGGTGTTGCTCTG 1217
QY 366 -----LeuProTyrLeuProSerGlyThrProGluLeuGlu 377
DB 1218 TGGCACTGTGCTTACGCTGTAGGTCTTCCATACTTGCCTTCAGGCACCCAGAGCTTGAA 1277
QY 378 ValProLysProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThr 397
DB 1278 GTTCCAAACACAGGTACGCCCTCTGAACCTGAGCATGGCCCAAGCCCTGGACAGACAG 1337
QY 398 PheCysGlnGlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyr 417
DB 1338 TTCTGCCAGGGCAAGCTGTAGTGGCTCTATCCCAATCTCTCGGAACGCTCCAGCTTCTAC 1397
QY 418 SerCysAlaAlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerAsn 437
DB 1398 AGCTGTGCAGCGGGGGCGCTGTTCAGCAAAAGCTGCCCGACAGGCGCTGTGTTTCAGCAAC 1457
QY 438 SerCysLysCysCysThrTrpAsn 445
DB 1458 TCCTGCAAAATGCTGCACCTGGAAT 1481
RESULT 11
US-09-151-011-5
; Sequence 5, Application US/09151011
; Patent No. 6057142
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: A Human Chitinase, Its Recombinant
; TITLE OF INVENTION: Production, Its Use For Decomposing Chitin, Its Use in
; TITLE OF INVENTION: Therapy or Prophylaxis Against Infection Diseases.
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; STREET: 6900 Jericho Turnpike
; CITY: Syosset
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11791
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; SOFTWARE:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/151.011
; FILING DATE: 10 - September - 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Morris, Robert C.
; REGISTRATION NUMBER: 42,910
; REFERENCE/DOCKET NUMBER: 294-32 DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1713 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; US-09-151-011-5
Alignment Scores:
Pred. No.: 5, 04e-233 Length: 1713
Score: 2348.50 Matches: 443
Percent Similarity: 94.66% Conservative: 0
Best Local Similarity: 94.66% Mismatches: 1
Query Match: 97.94% Indels: 25

DB: 3 Gaps: 1

US-10-004-219B-10 (1-445) x US-09-151-011-5 (1-1713)

QY 2 LysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPheLeu 21

DB 79 AAACCTGCTGCTACTTACCAACTGGGCCAGTACAGACAGGGGAGGCTCGCTTCTCTG 138

QY 22 ProLysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMetThr 41

DB 139 CCAAGGACTTGGACCCGACCCCTTGGACCCACCTCATCTACGCCCTTCGCTGGCATGACC 198

QY 42 AsnHisGlnLeuSerThrThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsnGly 61

DB 199 AACCCACCGCTGAGCACCACCTGAGTGGATGACGAGACTCTCTACAGGAGTTCAATGGC 258

QY 62 LeuLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPheGly 81

DB 259 CTGAAGAAGATGAATCCCAAGCTGAAGACCCCTGTAGCCATCGAGGCTGGAAATTCGGC 318

QY 82 ThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsnSer 101

DB 319 ACTCAGAAGTTTCAGATATGGTAGCCACGCGCCCAACACGTCAGACCTTTGTCAACTCG 378

QY 102 AlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyrPro 121

DB 379 GCCATCAGGTTTCTGCGCAATACAGCTTTGACGGCCTTGACCTTGGAGTAGTACCCA 438

QY 122 GlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAspLeu 141

DB 439 GGAAGCCAGGGAGCCCTGCGGTAGACAAGAGCGCTTCAACACCTGGTGACGAGCTTG 498

QY 142 AlaAsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuSerAla 161

DB 499 GCCAATGCCCTTCAGCAGGAAGCCAGACCTCAGGAAGACGCTTCTTCTGAGTGCA 558

QY 162 AlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAlaGln 181

DB 559 GCGGTTCCAGCTGGCAGACCTATGTGGATGCTGGATACGAGGTGACAAATCGCCAG 618

QY 182 AsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysValThr 201

DB 619 AACCTGGATTTTGTCAACCTTATGGCTACGACTTCCATGGCTTCTGGGAGAGGTCACG 678

QY 202 GlyHisAsnSerProLeuTyrLysArgGlnGlnLysSerGlyAlaAlaLaserLeuAsn 221

DB 679 GGAATTAACAGCCCTCTCAAGAGGCAAGAGAGTGCTGCAGCGCCAGCCTCAAC 738

QY 222 ValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuLeu 241

DB 739 GTGGATGCTGCTGTGCAACAGTGGCTGCAGAGGGGACCCCTGCCAGCAAGCTGATCCTT 798

QY 242 GlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgValGly 261

DB 799 GGCATGCTTACCTACGAGCCTCTTACACTGGCCTCTCTCATCAGACACAGAGTGGG 858

QY 262 AlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyGlyMetLeuAla 281

DB 859 GCCCCAGCCACAGGCTGGCATCTCAGGCCCTTCCACCAAGGAAGAGGATGCTGGCC 918

QY 282 TyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLysVal 301

DB 919 TACTATGAGTCTGCTCTGGAAGGGGGCCACCAACAGAGAATCCAGGATCAGAAGGTG 978

QY 302 ProTyrIlePheArgAspAsnGlnTrpValGlyPheAspValGluSerPheLysThr 321

DB 979 CCTATCATCTTCCGGGACCAACCAAGTGGGTGGCTTGTATGATGTGAGAGCTTCAAAACC 1038

QY 322 LysValSerTyrLeuLysGlnLysGlyLeuGlyGlyAlaMetValTrpAlaLeuAspLeu 341

DB 1039 AAGGTAGCTATCTGAAGCAGAGAGGACTGGGGGGGCCCATGCTTGGGCACTGGACTTA 1098

QY 342 AspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeuArg 361

DB 1099 GATGACTTTCGCGGCTTCTCTGCAACAGGGCCGATATACCCCTCATCCAGACGCTACGG 1158

QY 362 GlnGluLeuSer-----LeuProTyrLeuProSerGlyThrProGluLeuGlu 377

DB 1159 CAGGAACGTAA--TGGGTAAAGCCTTAACCTGCTGTCACTGCTGAGGACCCAGAGCTTGA 1217

QY 366 -----LeuProTyrLeuProSerGlyThrProGluLeuGlu 1277

DB 1218 TGGCACTGTGCTTACAGTGTAGTCTTCCATACCTTCCAGGACCCAGAGCTTGA 377

QY 378 ValProLysProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThr 397

DB 1278 GTTCCAAAACAGGTCAGCCCTCTGAACCTGAGCATGGCCCAAGCCCTGGACAGACAGC 1337

QY 398 PheCysGlnGlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerPheTyr 417

DB 1338 TTCTGCCAGGGCAAGCTGATGGGCTCTATCCCAATCTCGGGAACGGTCCAGCTTCTAC 1397

QY 418 SerCysAlaAlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerAsn 437

DB 1398 AGCTGTGACGGGGGGGCTGTTCAGCAAGCTGCCGACAGGCTGGTGTTCAGCAAC 1457

QY 438 SerCysLysCysCysThrTrpAsn 445

DB 1458 TCCTGCAATGCTGCACCTGGAAT 1481

RESULT 12

US-09-343-623-5

; Sequence 5, Application US/09343623

; Patent No. 6303118

; GENERAL INFORMATION:

; APPLICANT:

; TITLE OF INVENTION: A human chitinase, its recombinant

; TITLE OF INVENTION: production, its use for decomposing chitin, its use

; NUMBER OF INVENTIONS: 16

; CORRESPONDENCE ADDRESS:

; ADDRESSER: Hoffmann & Baron

; STREET: 350 Jericho Turnpike

; CITY: Jericho

; STATE: New York

; COUNTRY: United States of America

; ZIP: 11758

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; SOFTWARE:

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/343,623

; FILING DATE:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/486,839

; FILING DATE: 07-June-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Baron, Ronald J.

; REGISTRATION NUMBER: 29,281

; REFERENCE/DOCKET NUMBER: 294-26

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (516) 822-3550

; TELEFAX: (516) 822-3582

; INFORMATION FOR SEQ ID NO: 5:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1713 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: unknown

; TOPOLOGY: unknown

; MOLECULE TYPE: cDNA

; HYPOTHETICAL: NO

US-09-343-623-5

Alignment Scores:

Pred. No.: 5,04e-233 Length: 1713
Score: 2348.50 Matches: 443
Percent Similarity: 94.66% Conservative: 0
Best Local Similarity: 94.66% Mismatches: 1
Query Match: 97.94% Indels: 25
DB: 4 Gaps: 1

US-10-004-219B-10 (1-445) x US-09-343-623-5 (1-1713)

```
QY 2 LysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPheLeu 21
DB 79 AAACCTGGCTGCTACTTACCAACCTGGGCCAGTACAGACAGGGGAGGCTCGCTTCCTG 138
QY 22 ProLysAspLeuAspProSerLeuCysThrHisLeuLeuTyrAlaPheAlaGlyMetThr 41
DB 139 CCCAAGGACTTGACCCAGCCCTTTGCACCCACCTCATCTACGCCCTTCCTGGCATGACC 198
QY 42 AsnHisGlnLeuSerThrThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsnGly 61
DB 199 AACCAACAGCTCAGCACCACTCAGTGGATGACGAGACTCTCTACCAAGAGTTCAATGGC 258
QY 62 LeuLysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTyrAsnPheGly 81
DB 259 CTGAAGAGATGAATCCCAAGCTGAAGACCCTGTTAGCCATCGAGGCTGGAAATTCGGC 318
QY 82 ThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsnSer 101
DB 319 ACTCAGAAGTTTACAGATATGTAGTCCAGCGGCAACACCGTCAGACCTTTTGTCAACTCG 378
QY 102 AlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyrPro 121
DB 379 GCATCAGGTTTCTGCGCAATATACAGCTTTGACGGCTTGACCTGGAGTACCCA 438
QY 122 GlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAspLeu 141
DB 439 GGAAGCCAGGGAGCCCTCGCTAGCAAGAGGCGCTTCACACCCCTGTGTACAGACTTG 498
QY 142 AlaAsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuSerAla 161
DB 499 GCAATGCTTTCAGCAGAGCCAGACTCAGGAGAGAGCCCTTCTTCTGTAGTGA 558
QY 162 AlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAlaGln 181
DB 559 GCGGTTCCAGCTGGGCAGACCTATGTGGATGCTGGATACAGAGTGACAAAATCGCCAG 618
QY 182 AsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysValThr 201
DB 619 AACCTGGATTTTGTCAACCTTATGGCTACGACTTCATGGCTCTTGGGAGAGGTACG 678
QY 202 GlyHisAsnSerProLeuTyrLysArgGlnGluSerGlyValAlaAlaSerLeuAsn 221
DB 679 GGACATAACAGCCCTCTTACAGAGCGCAAGAGAGTGGTGACAGCCAGCCTCAAC 738
QY 222 ValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuIleLeu 241
DB 739 GTGGATGCTGCTGTCAACAGTGGCTGCAGAGGGGAGCCCTGCCAGCAAGCTGATCTT 798
QY 242 GlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgValGly 261
DB 799 GGCATGCCCTACCTACGAGGCTCTTTCACATGGCTCCCTCATCAGACCAAGAGTGGG 858
QY 262 AlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyGlyMetLeuAla 281
DB 859 GCCCCAGCCAGAGGCTCTGCACTCCAGGCCCTTACCAGAGAGGAGGATGCTGGCC 918
QY 282 TyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLysVal 301
DB 919 TACTATGAGTCTGCTCTCTGGAGGGGGGCCACCAACAGAGATCAGAGTCAAGAGGTG 978
QY 302 ProTyrIlePheArgAspAsnGlnTrpValGlyPheAspValGluSerPheLysThr 321
DB 979 CCTACATCTTCCGGGACCAACAGTGGGTGGGCTTTGATGATGTGGAGAGCTTCAAAACC 1038
```

```
QY 322 LysValSerTyrLeuLysGlnLysGlyLeuGlyGlyAlaMetValTrpAlaLeuAspLeu 341
DB 1039 AAGGTACGTATCTGAGACAGAGGACTGGCGGGCCATGGTCTGGGCACTGACITTA 1098
QY 342 AspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeuArg 361
DB 1099 GATGACTTTGGCGGCTTCTCTGCAACACGAGGCGGATACCCCTCATCCAGACGCTACGG 1158
QY 362 GlnGluLeuSer----- 365
DB 1159 CAGGAACCTGAA-TGGGTAAAGCCTTAACGTGCTGCATGTAGGCCAGGTTGTCCTG 1217
QY 366 -----LeuProTyrLeuProSerGlyThrProGluLeuGlu 377
DB 1218 TGGCAGCTGCTTTCAGCTGTAGTCTTCCATACCTTCAGGCCACCCAGAGCTTGAA 1277
QY 378 ValProLysProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThr 397
DB 1278 GTTCCAAACACGAGTCAAGCTCTGAACTGAGCATGGCCCAAGCCCTGGACAAGACAGC 1337
QY 398 PheCysGlnGlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyr 417
DB 1338 TTCTGCCAGGGCAAGCTGTATGGGCTCTATCCCAATCTCGGAAACGGTCCAGCTTCTAC 1397
QY 418 SerCysAlaAlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerAsn 437
DB 1398 AGCTGTGAGCGGGGCGGCTGTTCAGCAAGAGCTGCCCGCAGAGGCTGGTGTTCAGCAC 1457
QY 438 SerCysLysCysCysThrTrpAsn 445
DB 1458 TCCTGCAAAATGTCACCTGGAAT 1481
RESULT 13
US-08-694-915-1
; Sequence 1, Application US/08694915
; Patent No. 5811535
; GENERAL INFORMATION:
; APPLICANT: Adamou, Julie
; APPLICANT: Kirkpatrick, Robert.
; APPLICANT: Rosenberg, Martin
; TITLE OF INVENTION: HUMAN CARTILAGE GP39-LIKE GENE
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSER: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-2799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/694,915
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Han, William T
; REGISTRATION NUMBER: 34,344
; REFERENCE/DOCKET NUMBER: ATG50017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5219
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1433 base pairs
```

10

TELECOMMUNICATION INFORMATION:

TELEPHONE: 619/455-5100
TELEFAX: 619/455-5110
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1681 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
IMMEDIATE SOURCE:
CLONE: YKL-40
FEATURE:
NAME/KEY: CDS
LOCATION: 135..1681
PCT-US94-07754-4

Alignment Scores:

Pred. No.: 8.69e-97 Length: 1681
Score: 1030.00 Matches: 202
Percent Similarity: 67.33% Conservative: 68
Best Local Similarity: 50.37% Mismatches: 115
Query Match: 42.95% Indels: 16
DB: 5 Gaps: 8

US-10-004-219B-10 (1-445) x PCT-US94-07754-4 (1-1681)

```
QY      2  LysLeuValCysThrPheThrAsnTTPAlaGlnTyrArgGlnGlyGluAlaArgPheLeu 21
DB      138 AAAGTGTCTGTCTACTACACCACTGCTGCTCCAGTACCGGAAGCGCATGGAGCTGCTC 197
QY      22 ProLysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMetThr 41
DB      198 CCAGATGCCCTTGACCGCTTCTCTGTGTACCCACATCATCTACAGCTTGCCAAATATAGC 257
QY      42 AsnHisGlnLeuSerThrThrGluTTPAsnAspGluThrLeuTyrGlnGluPheAsnGly 61
DB      258 AACGATCATCATGCACACCTCGGAGTGGATGATGTGACGCTCTACGCGCATGCTCAACACA 317
QY      62 LeuLysLysMetAsnProLysLeuLysThrLeuAlaIleGlyGlyTTPAsnPheGly 81
DB      318 CTCAACACACGAAACCCAACTCCAGAGACTCTCTGTCTCGGAGATGGAATTTGGG 377
QY      82 ThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsnSer 101
DB      378 TCTCAAGATTTTCCAAGATAGCTTCCACACCCAGAGTCGCCGGACTTTCATCAAGTCA 437
QY      102 AlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyrPro 121
DB      438 GTACCGCCATTCTTGGCGACCCATGGCTTTGTATGGCGGTGACCTTGCTGGCTCTACCT 497
QY      122 GlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAspLeu 141
DB      498 GGACGGAGA-----GACAAACACCATTTTACCACTTACCGCCCTAATCAAGGAATG 542
QY      142 AlaAsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuSerAla 161
DB      543 AAGCGCGAATTTATAAGGAAGCCAG---CCAGGGAAGAAAGCAGCTCTCTGCTCAGCGCA 599
QY      162 AlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAlaGln 181
DB      600 GCAGTGTCTCGGGGAAGGTTCACCATTCAGCAGCGTATGACATTGCCAAGATATCCCAA 659
QY      182 AsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysValThr 201
DB      660 CACTGTGATTTCATAGCATGACCTACGATTTCATGGCGCTTGGCGTGGGACCACA 719
QY      202 GlyHisAsnSerProLeuTyrLysArgGlnGlnGluSerGlyAlaAlaAsnLeuAsn 221
DB      720 GGCCATCACAGTCCCTGTTCCGAGGTCCAGGAGATGCAAGTCTCCACAGATTCCAGCAAC 779
QY      222 ValAspAlaValGlnGlnThrProLysGlyThrGlnLysGlyThrProLysLysLeuIleLeu 241
```

Search completed: June 30, 2003, 01:11:43

Job time : 89.5244 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: June 29, 2003, 23:54:39 ; Search time 162.267 Seconds
(without alignments)
4070.941 Million cell updates/sec

Title: US-10-004-219B-10
Perfect score: 2398
Sequence: 1 AKLVCFYTNWQYRQGEARF.....QQSCPTGLVFNCKCCTWN 445

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1055720 seqs, 742224136 residues

Total number of hits satisfying chosen parameters: 2111440

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Command line parameters: -MODEL=frame+ p2n.model -DEV=xl
-Q/cgn2_1/USFTO_spool/US10004219/runat_24062003_160230_8285/app_query.fasta_1.1877
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10004219 @CGN 1 1 123 @runat_24062003_160230_8285
-NCPUS=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPHLOC=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/prodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/prodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/prodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/prodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/prodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/prodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq.*
10: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/prodata/1/pubpna/US10_NEW_PUB.seq.*
12: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/prodata/1/pubpna/US60_NEW_PUB.seq.*
14: /cgn2_6/prodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	1323	55.2	1625	9 US-10-004-219B-2 Sequence 2, Appli
2	1301.5	54.3	1525	9 US-10-004-219B-3 Sequence 3, Appli
3	1051	43.8	1925	9 US-10-097-340-44 Sequence 44, Appli
4	1025	42.7	1681	10 US-09-215-077A-4 Sequence 4, Appli

5	1025	42.7	1681	10	US-09-262-213A-4	Sequence 4, Appli
6	996	41.5	1391	10	US-09-822-830A-402	Sequence 402, App
7	951	39.7	1474	10	US-09-765-231A-8	Sequence 8, Appli
8	762	31.8	1470	9	US-10-218-743-40	Sequence 40, Appl
9	762	31.8	1470	9	US-10-218-743-42	Sequence 42, Appl
10	762	31.8	1527	9	US-10-218-743-37	Sequence 37, Appl
11	762	31.8	1527	9	US-10-218-743-39	Sequence 39, Appl
12	762	31.8	1621	9	US-10-218-743-34	Sequence 34, Appl
13	762	31.8	1621	9	US-10-218-743-36	Sequence 36, Appl
14	738	30.8	1608	9	US-10-218-743-20	Sequence 20, Appl
15	738	30.8	1608	9	US-10-218-743-22	Sequence 22, Appl
16	738	30.8	1665	9	US-10-218-743-17	Sequence 17, Appl
17	738	30.8	1665	9	US-10-218-743-19	Sequence 19, Appl
18	738	30.8	1752	9	US-10-218-743-14	Sequence 14, Appl
19	738	30.8	1752	9	US-10-218-743-16	Sequence 16, Appl
20	536	22.4	1140	9	US-09-938-842A-1684	Sequence 1684, Ap
21	524.5	21.9	1038	9	US-10-202-436A-14	Sequence 14, Appl
22	521.5	21.7	1797	10	US-09-974-300-646	Sequence 646, App
23	499	20.8	449	10	US-09-960-352-678	Sequence 678, App
24	487	20.3	1271	9	US-09-923-844B-1	Sequence 1, Appli
25	411	17.1	427	10	US-09-960-352-3315	Sequence 3315, Ap
26	410	17.1	438	10	US-09-960-352-1882	Sequence 1882, Ap
27	409	17.1	422	10	US-09-960-352-5957	Sequence 5957, Ap
28	409	17.1	423	10	US-09-960-352-12100	Sequence 12100, A
29	409	17.1	430	10	US-09-960-352-1189	Sequence 1189, Ap
30	404	16.8	426	10	US-09-960-352-4979	Sequence 4979, Ap
31	403	16.8	419	10	US-09-960-352-10172	Sequence 10172, A
32	402	16.8	422	10	US-09-960-352-10642	Sequence 10642, A
33	401	16.7	427	10	US-09-960-352-4978	Sequence 4978, Ap
34	399	16.6	419	10	US-09-960-352-4053	Sequence 4053, Ap
35	395	16.5	410	10	US-09-960-352-8317	Sequence 8317, Ap
36	394	16.4	429	10	US-09-960-352-14230	Sequence 14230, A
37	392	16.3	410	10	US-09-960-352-10161	Sequence 10161, A
38	390	16.3	440	10	US-09-960-352-4058	Sequence 4058, Ap
39	386	16.1	402	10	US-09-960-352-11601	Sequence 11601, A
40	384	16.0	424	10	US-09-960-352-3057	Sequence 3057, Ap
41	380	15.8	410	10	US-09-960-352-5496	Sequence 5496, Ap
42	378	15.8	437	9	US-09-918-995-35463	Sequence 35463, A
43	377	15.7	414	10	US-09-960-352-256	Sequence 256, App
44	374	15.6	399	10	US-09-960-352-265	Sequence 265, App
45	373	15.6	410	10	US-09-960-352-4536	Sequence 4536, Ap

ALIGNMENTS

RESULT 1
US-10-004-219B-2
; Sequence 2, Application US/10004219B
; Publication No. US20030087414A1
; GENERAL INFORMATION:
; APPLICANT: Macrozyme
; APPLICANT: Aerts, Johannes M.F.G.
; APPLICANT: Boot, Rolf G.
; TITLE OF INVENTION: A mammalian mucinase, its recombinant production, and its use in therapy or prophylaxis against diseases in mammals
; TITLE OF INVENTION: its use in therapy or prophylaxis against diseases in mammals
; TITLE OF INVENTION: which mucus is involved or infection diseases
; FILE REFERENCE: 2183-5136US
; CURRENT APPLICATION NUMBER: US/10/004,219B
; CURRENT FILING DATE: 2001-11-02
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 2
; TYPE: DNA
; LENGTH: 1625
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Combined DNA/RNA Molecule: human AMCase cDNA sequence and deduced amino acid sequence
; OTHER INFORMATION: AMCase cDNA sequence and deduced amino acid sequence
; OTHER INFORMATION: sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: human AMCase cDNA sequence and deduced amino acid sequence
; OTHER INFORMATION: cDNA sequence and deduced amino acid sequence
; FEATURE:

US-10-004-219B-10 (1-445) x US-10-004-219B-3 (1-1525)

Qy	3	LeuValCysTyrPheThrAenTrrpAlaGlnTyrArgGlnGlyGlualaaArgPheLeuPro	22
Db	70	CTGATATGCTATTTCACCAACTGGCCAGATATCGCCAGGTCTGGGGAGCTTCAAGCCT	129
Qy	23	LysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMetThrAsn	42
Db	130	GATGACATTAACCCCTGCTGTGTACTCACCTGATCTATGCCCTTGTGGATGCAGAAC	189
Qy	43	HisGlnLeuSerThrThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsnGlyLeu	62
Db	190	AATGAGATCACCAACCATAGAAATGAATGTTTACTCTCTATAAAGCTTTCAATGACTTG	249
Qy	63	LysLysMetAenProLysLeuLysThrLeuLeuAlaIleGlyGlyTyrAsnPheGlyThr	82
Db	250	AAAACAGGAACACCAAACTGAAACCCCTCTGGCAATTGGAGCTGGAACTTTGGAACT	309
Qy	83	GlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsnSerAla	102
Db	310	GCTCCTTTCACTACCATGGTTTCCACTTCTCAGAACCCGACAGCTTTCATTAACCTCAGTC	369
Qy	103	IleArgPheIleuArgLysTyrSerPheAspGlyLeuAspLeuAspTrrpGluTyrProGly	122
Db	370	ATCAAAATTTCTGCTGCTAGTATGGGTGTTGATGACTGGACCTGGAAATACCCAGGC	429
Qy	123	SerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAspLeuAla	142
Db	430	TCAGTGGGAGCCCTCCTCAGGACACAGCATCTCTTCACTGCTGCTGGTGAAGAAATCGGT	489
Qy	143	AsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuLeuSerAlaAla	162
Db	490	GAAGCTTTTTCAGCAGGAGGCTATTTCAGAGCAACAGAGCCGACACTGATGGTTACTTGCTGTCT	549
Qy	163	ValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAlaGlnAsn	182
Db	550	GTACTGGTGGGATTTCAACATCCAGCTGGCTATGAGATCCCTGAACTTTCCTTAAGTAC	609
Qy	183	LeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrrpGluLysValThrGly	202
Db	610	CTGGATTTTCATCCATGTGACATATGACCTCCATGGCTCCTGGAGGGGTACACTGGG	669
Qy	203	HisAsnSerProLeuTyrLysArgGlnGluSerGlyAlaAlaAlaSerLeuAsnVal	222
Db	670	GAGAAATAGTCTCTTTACAAATACCTACTGAGACTGGTAGCAATGCTTACCTCAATGTG	729
Qy	223	AspAlaIaValGlnGlnTrrpLeuGlnLysGlyThrProAlaSerLysLeuIleLeuGly	242
Db	730	GATTATGTCATGAATATTGGAAGAACAAATGGAGCCCGAGCTGAGAGCTCATTTGTGGA	789
Qy	243	MetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgValGlyAla	262
Db	790	TTCCACAGATATGGACACACCTTCATCTCGAAGAACCCCTCTGATAATGGAATTTGGTGCC	849
Qy	263	ProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyGlyMetLeuAlaTyr	282
Db	850	CCTACCTCTGGTGATGGCCCTGCTGGCGCCCTATACACAGAGGCTGGGTTCTGGGCGCTAC	909
Qy	283	TyrGluValCysSerTrrp-----LysGlyAlaThrLysGlnArgIleGlnAspGlnLys	300
Db	910	TATGAGATTTGCACTTTCTCAGAAAGTGGAGCCACTGAGGTCTGGGATGCCCTCCCAAGAA	969
Qy	301	ValProTrrpIlePheArgAspAsnGlnTrrpValGlyPheAspAspValGluSerPheLys	320
Db	970	GTGGCCCTATGCTTATAAGGCCAACAGAGTGCGTCTGGCTATGACATATCAAGAGCTTCAGT	1029
Qy	321	ThrLysValSerTrrpLeuLysGlnLysGlyLeuGlyGlyAlaMetValTrrpAlaLeuAsp	340
Db	1030	GTTAAGGCTCAGTGGCTTAACGACAGAACAAATTTTGAGGTGCCATGATCTGGGCCCATTTGAC	1089
Qy	341	LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeu	360
Db	1090	CTTGATGACTTTCATCGGCTCTTTCTGTGATCAGGAGAAATTTCTCTGACTTCTTACTTTG	1149

[illegible]

RESULT 3

US-10-097-340-44
; Sequence 44, Application US/10097340
; Publication No. US20030087250A1

```

: GENERAL INFORMATION:
: APPLICANT: John MONAHAN
: APPLICANT: Manjula GANNAVARAPU
: APPLICANT: SEBASTIAN HOERSCH
: APPLICANT: Shubhangi KAMATKAR
: APPLICANT: Steve G. KOVATS
: APPLICANT: Rachel E. MEYERS
: APPLICANT: Michael MORRISBY
: APPLICANT: Peter OLANDT
: APPLICANT: Ami SEN
: APPLICANT: Peter VIBY
: APPLICANT: Gordon B. MILLS
: APPLICANT: Robert C. EAST, Jr.
: APPLICANT: Karen LU
: APPLICANT: Rosemarie SCHMANDT
: APPLICANT: Xumei ZHAO
: APPLICANT: Karen GIATT

```

Alignment Scores:		
Pred. No.:	2.12e-105	Length:
Score:	1051.00	Matches:
Percent Similarity:	68.56%	Conservative:
		70
		1925

Best Local Similarity: 51.24% Mismatches: 107
Query Match: 43.83% Indels: 20
DB: 9 Gaps: 10

US-10-004-219B-10 (1-445) x US-10-097-340-44 (1-1925)

```
QY 2 LysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPheLeu 21
DB 193 AAACCTGGTCTGCTACTACACAGCTGGTCCAGTACCGGGAAGCGATGGAGCTGCTTC 252
QY 22 ProLysAspLeuAspProSerLeuCysThrHisLeuLeuTyrAlaPheAlaGlyMetThr 41
DB 253 CCAGATGCCCTTGCACCGCTCTCTGTACCCACATCATCTACAGCTTTGCCAATAAAGC 312
QY 42 AsnHisGlnLeuSerThrThrGluTyrAsnAspGluThrLeuTyrGlnGluPheAsnGly 61
DB 313 AACGATCATCATCGACACCTGGGAGTGGATGATGTACGCTCTACGGCATGCTCAACACA 372
QY 62 LeuLysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyTrpAsnPheGly 81
DB 373 CTCACGAACACAGGAACCCCAACCTGAAGACTCTCTGTCTGTGGAGGATGGAACTTTGGG 432
QY 82 ThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsnSer 101
DB 433 TCTCAAGATTTTCCAAAGATAGCTCCCAACACACAGAGTGGCGGACTTTTCATCAAGTCA 492
QY 102 AlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyrPro 121
DB 493 GTACCCGCAATCTCTGGCACCACCATGCTTTGATGGGCTGGACCTTCCTGGCTTACCT 552
QY 122 GlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAspLeu 141
DB 553 GGACGGGAGA-----GACAAACAGCATTTTACCACCCTAATCAAGGAATG 597
QY 142 AlaAsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuSerAla 161
DB 598 AAGGCCGAATTTTAAAGAAAGCCAG---CCAGGGAAGAAAGCAGCTCTCTGCTACGCGCA 654
QY 162 AlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAlaGln 181
DB 655 GCATCTCTCGGGGAGGTCCACCATGACAGCATATGACATTCACCATGATATCCCAA 714
QY 182 AsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysValThr 201
DB 715 CACCTGGATTTCATTAGCATGACTACGATTTTCATGGAGCTGGCGTGGGACCA 774
QY 202 GlyHisAsnSerProLeuTyrLysArgGlnGluSerGlyAlaAlaSerLeuAsn 221
DB 775 GGCCATCAGATGCTCCCTGTTCGAGTCAAGGAGTGCAGAGTCCCTGACAGATTCAGCAAC 834
QY 222 ValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuLeu 241
DB 835 ACTGACTATGCTGTGGGTACATGTTGAGGCTGGGCTCTCTGCCAGTAACTGGTGATG 894
QY 242 GlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgValGly 261
DB 895 GGATCCCAACCTTCGGAGGAGCTTCACTGTGCT---TCTTCTGACAGCTGGTGTGA 951
QY 262 AlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyGlyMetLeuAla 281
DB 952 GCGCCCAATCTCAGGACCGGGAATTCAGGCGCGGTTCACCAAGAGGACGCGGACCTTGGC 1011
QY 282 TyrTyrGluValCysSerTrp---LysGlyAlaThrLysGlnArgIleGlnAspGlnLys 300
DB 1012 TACTATGATCTGTGACTCTCTCCGGGAGGCACAGTCCATAGAACCTTCGGCCAGAG 1071
QY 301 ValProTyrIlePheAspAsnGlnTrpValGlyPheAspAspValGluSerPheLys 320
DB 1072 GTCCCTATGCCAACCAAGCAACCAAGTGGTAGGATACGACGACGAGGAAGCGTCAAA 1131
QY 321 ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyGlyAlaMetValTrpAlaLeuAsp 340
DB 1132 AGCAAGTGCAGTACCTGAAGGATAGGACGCTGGCAGGCGCCATGTATGGGCCCTGGAC 1191
```

```
QY 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGly---ArgTyrProLeuIleGlnThr 359
DB 1192 CTGGATGACTTCCAGGGCTCCTTCTGCGGCAGGATCTGGCGTCCCTCTCACAATGCC 1251
QY 360 LeuArgGlnGluLeu-----SerLeuProTyrLeuProSerGlyThrPro 374
DB 1252 ATCAAGGATGCACTCGCTGCAACAGTAGCCCTCTGTGTGCAACACAGCAGCGGG---CCA 1308
QY 375 GluLeuGluValProLysProGlyGlnProSer-----GluProGluHisGlyProSer 392
DB 1309 AGGATGCCCGCTGCC---CCTGTGCTCAGCTGGCGGAGCCTGATCACCTGCCCTGC 1365
QY 393 -----ProGly 394
DB 1366 TGAGTCCCAGGC 1377

RESULT 4
US-09-215-077A-4
; Sequence 4, Application US/09215077A
; Patent No. US20020031793A1
; GENERAL INFORMATION:
; APPLICANT: PRICE, PAUL A.
; APPLICANT: JOHANSEN, JULIA S.
; TITLE OF INVENTION: ASSAY FOR YKL-40 AS A MARKER FOR DEGRADATION OF
; FILE OF INVENTION: MAMMALIAN CONNECTIVE TISSUE MATRICES
; FILE REFERENCE: 407T-895411US
; CURRENT APPLICATION NUMBER: US/09/215,077A
; CURRENT FILING DATE: 1998-12-18
; PRIOR APPLICATION NUMBER: 08/581,527
; PRIOR FILING DATE: 1996-04-17
; PRIOR APPLICATION NUMBER: 08/089,989
; PRIOR FILING DATE: 1993-07-09
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1681
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: cDNA nucleotide
; OTHER INFORMATION: sequence for the coding region of the gene for
; OTHER INFORMATION: YKL-40.
US-09-215-077A-4

Alignment Scores:
Pred. No.: 1,256-102 Length: 1681
Score: 1025.00 Matches: 202
Percent Similarity: 67.08% Conservative: 67
Best Local Similarity: 50.37% Mismatches: 116
Query Match: 42.74% Indels: 16
DB: 10 Gaps: 8

US-10-004-219B-10 (1-445) x US-09-215-077A-4 (1-1681)
QY 2 LysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPheLeu 21
DB 138 AAACCTGGTCTGCTACTACACAGCTGGTCCCAAGTACCGGGAAGCGATGGAGCTGCTTC 197
QY 22 ProLysAspLeuAspProSerLeuCysThrHisLeuLeuTyrAlaPheAlaGlyMetThr 41
DB 198 CCAGATGCCCTTGCACCGCTTCTGTGTATCCCAACATCATCTACAGCTTTGCCAATAAAGC 257
QY 42 AsnHisGlnLeuSerThrThrGluTyrAsnAspGluThrLeuTyrGlnGluPheAsnGly 61
DB 258 AACGATCATCATCGACACCTGGGAGTGGATGATGTACGCTCTACGGCATGCTCAACACA 317
QY 62 LeuLysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyTrpAsnPheGly 81
DB 318 CTCACGAACACAGGAACCCCAACCTGAAGACTCTCTGTCTGTGGAGGATGGAACTTTGGG 377
QY 82 ThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsnSer 101
DB 1192 AGCAAGTGCAGTACCTGAAGGATAGGACGCTGGCAGGCGCCATGTATGGGCCCTGGAC 1191
```

378 TCTCAAGATTTCACAGTAGCTCCCAACCCAGAGTCGCGGACTTTCATCAAGTCA 437
QY 102 AlalileAArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyrPro 121
Db 438 GTACCGCCATTCTCGGACCCATCGCTTGTATGGCGTGACCTTGCTGGCTTACCT 497
QY 122 GlySerGlnGlySerProAlaValAspLysGluArgPheThrLeuValGlnAspLeu 141
Db 498 GGACCGGAGA-----GACAAACACCATTTTACCACCTAATCAAGAAATG 542
QY 142 AlaAsnAlaPheGlnGlnGlnThrSerGlyLysGluArgLeuLeuSerAla 161
Db 543 AAGGCCGAATTTATTAAGGAAGCCAG---CCAGGGAAGAACAGCTCTCTGCTCAGCGCA 599
QY 162 AlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAlaGln 181
Db 600 GCATGTCTGCGGGAAGGTTCACCATTCACAGCATGATGACATTCACCATGATATCCCAA 659
QY 182 AsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysValThr 201
Db 660 CACCTGGATTTTCATTAGCATCATGACCTACGATTTTATGCGCGCTGGCGTGGACCA 719
QY 202 GlyHisAsnSerProLeuTyrLysArgGlnGluSerGlyAlaAlaSerLeuAsn 221
Db 720 GGCCATCAGTCCCTCAGCGCGAGTCCAGGAGTCAAGTCTGACAGATTCAGCAAC 779
QY 222 ValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuLeu 241
Db 780 ACTGATATGCTGGGTGACATGTTGAGCTGGGGCTCTCCACGATGAGTGGTGG 839
QY 242 GlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerAspThrArgValGly 261
Db 840 GGCAATCCACCTCGGAGGAGCTTCACTCGCT---TCTTCTGAGACTGGTGTTC 896
QY 262 AlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyMetLeuAla 281
Db 897 GCGCAATCTCAGGACCGGGAATTCAGCGCGGTTCCACCAAGGAGCGAGCCCTG 956
QY 282 TyrTyrGluValCysSerTrp---LysGlyAlaThrLysGlnArgIleGlnAspGlnLys 300
Db 957 TACTATGATGCTGTGACTTCTCGCGGAGGACAGTCCATAGAACCTCGCGGAG 1016
QY 301 ValProTyrIlePheArgAspAsnGlnTrpValGlyPheAspValGluSerPheLys 320
Db 1017 GTCCCTATGTCACCAAGGCAACCACTGGGTAGGATACGACGACGAGAAAGCTCAA 1076
QY 321 ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyGlyAlaMetValTrpAlaLeuAsp 340
Db 1077 AGCAAGGTGCATCTGAAAGGATAGGCGAGCTGGGAGCGCCCATGATGGGCCCTGG 1136
QY 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGly---ArgTyrProLeuIleGlnThr 359
Db 1137 CTGGATGACTTCAGGGCTCTCTGCGGCGAGGATGCGGCTTCTCTCACCATG 1196
QY 360 LeuArgGlnGluSerLeuProTyrLeuProSerGlyThrProGluLeuGluValPro 379
Db 1197 ATCAAGGATGCACTCGCTCAAG---TAGCCCTCTGTTCTGCACACAGCAGCGGGCCA 1253
QY 380 LysProGlyGlnPro-----SerGluProGluHisGlyProSer-----Pro 393
Db 1254 AGGATGCCCGCTCCCGCTGCTGCGCGGAGCCTGATCACCTGCGCTGTGAGTCCCA 1313
QY 394 Gly 394
Db 1314 GGC 1316

RESULT 5

US-09-262-213A-4
; Sequence 4, Application US/09262213A
; Patent No. US20020090658A1
; GENERAL INFORMATION:
; APPLICANT: PRICE, PAUL
; APPLICANT: JOHANSEN, JULIA

; TITLE OF INVENTION: ASSAY FOR YKL-40 AS A MARKER FOR DEGRADATION OF
; FILE OF INVENTION: MAMMALIAN CONNECTIVE TISSUE MATRICES
; CURRENT APPLICATION NUMBER: US/09/262.213A
; CURRENT FILING DATE: 1999-03-04
; PRIOR APPLICATION NUMBER: US 08/089,989
; PRIOR FILING DATE: 1993-07-09
; PRIOR APPLICATION NUMBER: PCT/US94/07754
; PRIOR FILING DATE: 1994-07-08
; PRIOR APPLICATION NUMBER: US 08/581,527
; PRIOR FILING DATE: 1996-04-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 4
; LENGTH: 1681
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-262-213A-4
Alignment Scores:
Pred. No.: 1,25e-102 Length: 1681
Score: 1025.00 Matches: 202
Percent Similarity: 67.08% Conservative: 67
Best Local Similarity: 50.37% Mismatches: 116
Query Match: 42.74% Indels: 16
DB: 10 Gaps: 8
US-10-004-219B-10 (1-445) x US-09-262-213A-4 (1-1681)

QY 2 LysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPheLeu 21
Db 138 AAACCTGGTCTGTACTACACAGCTGGTCCAGTACCGGGAAGCGATGGGAGCTGCTTC 197
QY 22 ProLysAspLeuAspProSerLysCysThrHisLeuIleTyrAlaPheAlaGlyMetThr 41
Db 198 CCAGATGCCCTTGACCGCTTCTGTGTACCCACATCATCAGCTTTGCCAATAAAGC 257
QY 42 AsnHisGlnLeuSerThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsnGly 61
Db 258 AACGATCATCATCGACACCTGGGAGTGGGAATGATGTGACGCTCTACGGCATGCTCAACA 317
QY 62 LeuLysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPheGly 81
Db 318 CTCACAAACAGAACCCCAACCTGAAGACTCTCTTGTCTGTGGAGGATGGAACTTTGGG 377
QY 82 ThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsnSer 101
Db 378 TCTCAAGATTTTCAAGATAGCTTCCAAACCCCAAGAGTCCCGGACTTTCATCAAGTCA 437
QY 102 AlalileAArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyrPro 121
Db 438 GTACCGCCATTCTGCGCACCCATGGCTTGTATGGCGTGACCTTGCTGGCTTACCT 497
QY 122 GlySerGlnGlySerProAlaValAspLysGluArgPheThrLeuValGlnAspLeu 141
Db 498 GGACCGGAGA-----GACAAACACCATTTTACCACCTAATCAAGAAATG 542
QY 142 AlaAsnAlaPheGlnGlnGlnThrSerGlyLysGluArgLeuLeuSerAla 161
Db 543 AAGGCCGAATTTATTAAGGAAGCCAG---CCAGGGAAGAACAGCTCTCTGCTCAGCGCA 599
QY 162 AlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAlaGln 181
Db 600 GCATGTCTGCGGGAAGGTTCACCATTCACAGCATGATGACATTCACCATGATATCCCAA 659
QY 182 AsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysValThr 201
Db 660 CACCTGGATTTTCATTAGCATCATGACCTACGATTTTATGCGCGCTGGCGTGGACCA 719
QY 202 GlyHisAsnSerProLeuTyrLysArgGlnGluSerGlyAlaAlaSerLeuAsn 221
Db 720 GGCCATCAGTCCCTCAGCGCGAGTCCAGGAGTCAAGTCTGCTGACAGATTCAGCAAC 779

QY 319 PheLysValSerThrLeuLysGlnLysGlyLeuGlyAlaMetValTrrPala 338
Db 1021 TTGGGGCCCAAGTTTCAGTTCTTAAAGAAATTTAAACCTGGGGGTGCCTTGATTGGTCT 1080
QY 339 LeuAspLeuAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGln 358
Db 1081 TTTGACATGGTTGACTTCACTGGCAAACTCTGCAACCGGGGCCCTTCCCTCTTGTGCCAA 1140
QY 359 ThrLeuArgGlnGluLeu----- 364
Db 1141 GCAGTCAGAGAGAGCTTGGCTCCCTGTGAAGGATTAACCTACAGAGAGAGGCAAGAT 1200
QY 365 -----SerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLysPro 381
Db 1201 GCCCTTGTGCTGGGGCTGCTCTCTCCAGGAATTCATGTGGGATTCCTCCCTTGCCA 1260
QY 382 GlyGlnProSerGluProGluHisGlyProSerProGlyGlnAsp----- 396
Db 1261 GCCCGGCTTTGGATCTCTCTTCCAAGCCTTTCCTGACTTCTCTTAGATCATAGATTGG 1320
QY 397 ---ThrPheCysGlnGlyLysAlaAspGlyLeuTyrPro 408
Db 1321 ACCTGGTTTGGTTTCTTCGACGTCATGCCCTNTTGGCC 1359

RESULT 7

US-09-765-231A-8
; Sequence 8, Application US/09765231A
; Patent No. US20020119452A1
; GENERAL INFORMATION:
; APPLICANT: Searle/Monsanto
; APPLICANT: Phippard, Deborah
; APPLICANT: Vasanthakamur, Geetha
; APPLICANT: Dotson, Stanton
; APPLICANT: Ma, Xiao-Jun
; TITLE OF INVENTION: Osteoarthritis tissue-derived nucleic acids, polypeptides,
; FILE OF INVENTION: vectors, and cells
; FILE REFERENCE: SO-3221 PR
; CURRENT APPLICATION NUMBER: US/09/765,231A
; CURRENT FILING DATE: 2001-01-18
; NUMBER OF SEQ ID NOS: 82
; SEQ ID NO 8
; LENGTH: 1474
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-765-231A-8

Alignment Scores:

Pred. No.: 1.37e-94 Length: 1474
Score: 951.00 Matches: 196
Percent Similarity: 66.19% Conservative: 80
Best Local Similarity: 47.08% Mismatches: 108
Query Match: 39.68% Indels: 33
Gaps: 5

US-10-004-219B-10 (1-445) x US-09-765-231A-8 (1-1474)

QY 2 LysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPheLe 21
Db 140 AAACCTGGTTTGTACTTTTACCACCTGGGTCCAGGACCGGAGAACCAAGGAAATTCAC 199
QY 21 uPro-LysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPhe-AlaGlyMet 40
Db 200 CCTCAGGAATATTGACCCCTTCCATGCTCTCATCTCATCTATTCAATTCGCCAGCATC 259
QY 41 ThrAsnHisGlnLeuSerThrThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsn 60
Db 260 GAAACACACAGAGTTATCATCAAGGACACAGAGTGAAGTGATGCTCTACACAGACCATCAAC 319
QY 61 Gly-LeuLysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTrrAsnPh 80
Db 320 AGTTCTCAAAACCAAGAAATCCCAAAATCTTGTCTCCATGGAGGAGTCACTGT 379
QY 80 eGlyThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAs 100

Db 380 TGGTTCCAAAGGGTTCCACCCTATGGTGGATTCTTACATCAGGCTTGGAAATTCATTAA 439
QY 100 nSerAlaIleArgPheLeuArgLysTyrSerPheAsnGlyLeuAspLeuAspTrrGluTyr 120
Db 440 CTCATATCTCTGTTCTTGAGGAACCATTAACCTTTTGATGGACTGGATGAAGCTGATCTA 499
QY 120 rProGlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAs 140
Db 500 CCCAGATCAGAAA-----GAAACACTCATTTCTACTGTGCTGATTCTATGA 544
QY 140 pLeuAlaAsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuLeuSe 160
Db 545 GTTAGCAGAACCTTTCAGAGGACTTCACAAAATCCACCAAGGAAGGCTTCTCTTGAC 604
QY 160 rAla-AlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLeuLea 180
Db 605 TCGGGGGGTATCTCGAGGGAGGCAATATGATGTATTAACAGCTATCAAGTTGAGAACTGG 664
QY 180 laGlnAsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrrGluLys- 199
Db 665 CAAAGATCTGATTTTCATCAACCTCTGCTCTTGTACTTCCATGGGTCTTGGGAAAGC 724
QY 200 -----ValThrGlyHisAsnSerPro-LeuTyrLysArgGlnGluSerGlyAlaAla 217
Db 725 CCTTATCACTGGCCACAAACAGCCCTGCTGAGCAAGGGGTGGCAGGACAGAGGGCCAAGC 784
QY 218 AlaSerLeuAsnValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSer 237
Db 785 TCCTACTACAATGTGGAATATGCTGTGGGTACTGGATACATATAAGGGAATGCCATCAGAG 844
QY 238 LysLeuIleLeuGlyMetProThrTyr-GlyArgSerPheThrLeuAlaSerSerSerAs 257
Db 845 AAGTGTGTCATGGGATCCCCACATATGGGCACCTCTTACACTGGCC---TCTGCAGA 901
QY 257 pThrArgValGlyAlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGl 277
Db 902 AACCACTGGGGGCCCTGCTCTGGCCCTGGAGCTGCTGGACCCATCACAGAGTCTTC 961
QY 277 yGlyMetLeuAlaTyrTyrGluValCysSerTrp---LysGlyAlaThrLysGlnArgI 296
Db 962 AGGCTTCTCTGCTTATATGATGATCTGCCAGTTCTCTGAAAGGAGGCAAGATCACCGGCT 1021
QY 296 eGlnAspGlnLysValProTyrIlePheArgAspAsnGlnTrpValGlyPheAspVa 316
Db 1022 CCAGGATCAGCAGGTTCCCTACGAGTCAAGGGNACCAGTGGGTGGGTATGATGATGT 1081
QY 316 lGluSerPheLysThrLysValSerTyrLeuLysGlnLysGlyLeuGlyGlyAlaMetVa 336
Db 1082 GAAGAGTATGGAGACCAAGGTTTCAGTTCTTAAAGAAATTTAAACCTGGGAGGAGCATGAT 1141
QY 336 lTrrAlaLeuAspLeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLe 356
Db 1142 CTGGTCTATTACATGGATGATCTTCACTGGCAAAATCTCTGCAACCCAGGGCCCTTACCCTCT 1201
QY 356 uIleGlnThrLeuArgGlnGluLeu----- 364
Db 1202 TGTCCACAGCTCAAGAGAGCCCTTGGCTCCCTGTGAAGGATTAACCTACAGAGAGCAG 1261
QY 365 -----SerLeuProTyrLeuProSerGlyThrProGluLeuGluValPro 379
Db 1262 GCAAGATCACCTTGTGCTGGGGCTGCTCTCTCCAGGAATTCATGTGGGATTCCTC 1321
QY 379 oLysProGlyGlnProSerGluProGluHisGlyPro 391
Db 1322 CTTCAGGCGGGCTTTGGATCTCTCTTCCAAGCCT 1358

RESULT 8

US-10-219-743-40
; Sequence 40, Application US/10218743
; Publication No. US2003009679A1
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.

QY	170	ValAspAlaGlyTyrGluValAspIleLeuAlaGlnAsnLeuAspPheValIleLeuMet	189
DB	556	ATTGACGTAGCTTATGAGCTCAAGAAATTTGAACCAATTTGTCGATTGGAATGATGTCATG	615
QY	190	AlaTyrAspPheHisGlySerTrpGluLysValThrGlyHisSerProLeuTyrLys	209
DB	616	ACTTATGATTACCATGGCGGATGGGAATAATGTTTCGGCCATAATGCTCCGTTGTATAAA	675
QY	210	ArgGlnGluSerGlyAlaAlaAspSerLeuAsnValAspAlaAlaValGlnGln	228
DB	676	CGACCGGATGAACGGATGAATGACACCTTACTTCAATGTCACACTACACCTGCACTAT	735
QY	229	TrpLeuGlnLysGlyThrProAspSerLysLeuLeuLeuGlyMetProThrTyrGlyArg	248
DB	736	TATTTGAACAATGGCGCTACTCGAGACAACTGTTATGGGTGTTCCATTCTATGTCGT	795
QY	249	SerPheThrLeuAlaSerSerSerAspThrArgValGlyAlaProAlaThrGlySerGly	268
DB	796	GCTTGGAGCTCGAAGATCGAAGCAAAAGTCAAACTGGCGATCCGGCCAAAGCATGTCT	855
QY	269	ThrProGlyProPheThrLysGluGlyGlyMetLeuAlaTyrTyrGluValCysSerTrp	288
DB	856	CTCTCTGGTTTATTACTGGTGAAGAGTGTCTCTCATACATCGAATGTGTCACTTA	915
QY	289	LysGlyAlaThrLysGlnArgIleGln---AspGln-----LysValProTyrIlePhe	305
DB	916	TTCCAGAAAGAAGATGGCATATTCAATACGATGAATATTCAATGCTCCATACGGATAT	975
QY	306	ArgAspAsnGlnTrpValGlyPheAspValGluSerPheLysThrLysValSerTyr	325
DB	976	AATGATAAAATCTGGGTGGTTTACGATGATCTGGGTAGTATCATCATGCAAGTTGGC	1035
QY	326	LeuLysGlnLysGlyLeuGlyGlyAlaMetValTrpAlaLeuAspLeuAspPheAla	345
DB	1036	CTCAAGAATTCGGCGTCTCGCGGTATGATATGCTCATTTGGAAAAACGATGATTTCAA	1095
QY	346	GlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeuArgGlnGluLeuSer	365
DB	1096	GGT---CATTTGGGACCG--AAATATCCATTGTTGAACAAAGTTCAAAATATGATCAAT	1149
QY	366	-----LeuPro	367
DB	1150	GGTGATGAAAGAACTCTTTACGAATGCTTTTGGGCCCAAGTACAAACACCAACACCA	1209
QY	368	TyrLeuProSerGlyThrProGluLeuGluValProLysProGlyGlnProSerGluPro	387
DB	1210	ACCAACCGGTCAACTACT-----TCGACTACCAACCAACCGCTACCAACCGAT	1260
QY	388	GluHisGlyProSerProGlyGlnAspThrPheCysGlnGlyLys-----	402
DB	1261	AGCAACAGCGAAACCAACCAAAATACACTAGTATATTGATGGACATTTGATTAATGCTAT	1320
QY	403	AlaAspGlyLeuTyrProAsnProArgSerGluArgSerSerPheTyrSerCys-----	419
DB	1321	AAACAAGGTATCTTCCACTCAACTGATGTTTCATAAATATTTAGTTTGTGAATATATT	1380
QY	420	-----AlaAlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPhe	435
DB	1381	GCCACCAACCAACCGGTGGTGGGTGACATATTGATTGTTCCAAAGGAAGCTAGATGG	1440
QY	436	SerAsnSerCysLysCysCys	442
DB	1441	CACGCAACATTAATAAATTTGT	1461

RESULT 9
 US-10-218-743-42/c
 ; Sequence 42, Application US/10218743
 ; Publication No. US20030096779A1
 ; GENERAL INFORMATION:
 ; APPLICANT: McCall, Catherine A.
 ; APPLICANT: Hunter, Shirley Wu
 ; APPLICANT: Weber, Eric R.

RESULT 9
US-10-218-743-42/c
; Sequence 42, Application US/10218743
; Publication No. US20030096779A1
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Weber, Eric R.

; PRIOR APPLICATION NUMBER: US/09/292,225
 ; PRIOR FILING DATE: 1999-04-15
 ; PRIOR APPLICATION NUMBER: 60/098,909
 ; PRIOR FILING DATE: 1998-09-02
 ; PRIOR APPLICATION NUMBER: 60/085,295
 ; PRIOR FILING DATE: 1998-05-13
 ; PRIOR APPLICATION NUMBER: 60/098,565
 ; PRIOR FILING DATE: 1998-04-17
 ; PRIOR APPLICATION NUMBER: 09/062,013
 ; PRIOR FILING DATE: 1998-04-17
 ; NUMBER OF SEQ ID NOS: 49
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 37
 ; LENGTH: 1527
 ; TYPE: DNA
 ; ORGANISM: Dermatophagoides farinae
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(1527)
 ; US-10-218-743-37

Alignment Scores:

Pred. No.: 8,1e-74 Length: 1527
 Score: 762.00 Matches: 162
 Percent Similarity: 52.36% Conservative: 93
 Best Local Similarity: 33.26% Mismatches: 172
 Query Match: 31.78% Indels: 60
 DB: 9 Gaps: 14

US-10-004-219b-10 (1-445) x US-10-218-743-37 (1-1527)

QY 2 LysLeuValCysThrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPheLeu 21
 DB 100 CGAATCGATGTTATGTTGGAAACATGTCGGTTTATCATAAAGTTGATCCA---TACACA 156
 QY 22 ProLysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMetThr 41
 DB 157 ATTGAAGATATTGATCTCTTCAAAATGTTACTCATTTGATGTTATGTTTGTCTAAATCGAT 216
 QY 42 AsnHisGlnLeuSerThr-----ThrGluTyr 50
 DB 217 GAATACAAATACACATTCATCAAGTTTGTATCCATTTCAAGATGATTAACATACTCATGG 276
 QY 51 AsnAspGluThrLeuTyrGlnGluPheAsnGlyLeuLysLysMetAsnProLysLeuLys 70
 DB 277 ---GAAAAACACCGGTATGAACGTTTCAACAACTTGAGATTGAAGATCCAGAATTGACC 333
 QY 71 ThrLeuLeuAlaIleGlyGlyTyrAsnGlyThrGlnLysPheThrAspMetValAla 90
 DB 334 ACCATGATTTCAATGGGTGGTGGTATGAAGGTTTCAGAAAAATATTTCGGATATGGCAGCC 393
 QY 91 ThrAlaAsnAsnArgGlnThrPheValAsnSerAlaIleArgPheLeuArgLysTyrSer 110
 DB 394 AATCCAAATATCGTCAGCAATTTGTTCAATAGTTTGGACATTTTGGCAAGTAATACAA 453
 QY 111 PheAspGlyLeuAspLeuAspTyrGluTyrProGlySerGln---GlySerProAlaVal 129
 DB 454 TTCGATGGCCCTAGATTGGATTGGGAATATCTCGTATGATCCAGGTTAGGCAATCTAAATC 513
 QY 130 AspGlyGluArgPheThrThrLeuValGlnAspLeuAlaAsnAlaPheGlnGlnGluAla 149
 DB 514 GATAAACAACAACTATTAACTAGTTAGAGAACTTAAAGAGGCAATTTGAACCTTTTCGGC 573
 QY 150 GlnThrSerGlyLysGluArgLeuLeuSerAlaAlaValProAlaGlyGlnThrTyr 169
 DB 574 -----TACTTGTGTGACTGCGCGAGTATCCACCCGGTAAGATAA 612
 QY 170 ValAspAlaGlyTyrGluValAspLysIleAlaGlnAsnLeuAspPheValAsnLeuMet 189
 DB 613 ATTGAGCTAGCTTATGAGCTCAAGAAGATTGAACCAATTTGTTGATGGATGATGTCATG 672
 QY 190 AlaTyrAspPheHisGlySerThrGluLysValThrGlyHisAsnSerProLeuTyrLys 209

DB 673 ACTTATGATTACCATGGCGGATGGGAAATGTTTTCGGCCATAATGTCCTCGTGTATATAA 732
 QY 210 ArgGlnGluGluSerGlyAlaAlaSer---LeuAsnValAspAlaAlaValGlnGln 228
 DB 733 CGACCCGATGAACCGATCAATTGCACACTTACTTCAATGTCAACTACACCATGCACAT 792
 QY 229 TrpLeuGlnLysGlyThrProAlaSerLysLeuIleLeuGlyMetProThrTyrGlyArg 248
 DB 793 TATTTGAAACAATGGCGCTACTCGAGACAACTTGTATGGGTGTTCCTTCTATGTCGT 852
 QY 249 SerPheThrLeuAlaSerSerSerAspThrArgValGlyAlaProAlaThrGlySerGly 268
 DB 853 GCTTGGAGCATCGAAGATCGAAGCAAGTCAACTTGGCGATCCGCGCAAGAGCATGCT 912
 QY 269 ThrProGlyProPheThrLysGlyGlyMetLeuAlaTyrTyrGluValCysSerTrp 288
 DB 913 CCTCTGGTTTATTACTCTGGTGAAGAAGTGTCTCTCATACATCGAATTTGTGTCAGTTA 972
 QY 289 LysGlyAlaThrLysGlnArgIleGln---AspGln-----LysValProTyrIlePhe 305
 DB 973 TTCCAGAAAGAAAGATGCGATATTCAATACGATGAATATTACAATGCTCCATACGATAT 1032
 QY 306 ArgAspAsnGlnTrpValGlyPheAspAspValGluSerPheLysThrLysValSerTrp 325
 DB 1033 AATGATAAAATCTGGTGGTTTACGATGATCTGGCTAGTATATCATGCAAGTTGGCCTTT 1092
 QY 326 LeuLysGlnLysGlyLeuGlyGlyAlaMetValTrpAlaLeuAspLeuAspPheAla 345
 DB 1093 CTCAAAGAATTGGGCGTCTCTGGCGTTATGATATGTCATTTGGAAACGATGATTTCAA 1152
 QY 346 GlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeuArgGlnGluLeuSer 365
 DB 1153 GGT---CATTCGGGCGG---AAATATCAATGTTTGAACAAAGTTTCACAATATGATCAAT 1206
 QY 366 -----LeuPro 367
 DB 1207 GGTGATGAAGAAAGAACTCTTACGAATGCTTTTGGCGCCAAAGTACAAACCACCAACCA 1266
 QY 368 TyrLeuProSerGlyThrProGluLeuGluValProLysProGlyGlnProSerGluPro 387
 DB 1267 ACCACCCGTCACACTACT-----TCGACTACCACCAACCAACCGCTTACCACCCGAT 1317
 QY 388 GluHisGlyProSerProGlyGlnAspThrPheCysGlnGlyLys----- 402
 DB 1318 AGCACAAGCGAAACCAACCAAAATACACTACGTATATTGATGGACATTTGATTAATGCTAT 1377
 QY 403 AlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyrSerCys----- 419
 DB 1378 AAACAAGGTATCTTCCACATCCAACTGATTTCAATAATTTAGTTTGTGAATATATT 1437
 QY 420 -----AlaAlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPhe 435
 DB 1438 GCCACACCAACCGTGGTGGGTACATATTGATTTGTTCCAAAGAACTAGATGG 1497
 QY 436 SerAsnSerCysLysCysCys 442
 DB 1498 CACGCAACATTAAAAAATTGT 1518

RESULT 11

US-10-218-743-39/c
 ; Sequence 39, Application US/10218743
 ; Publication No. US20030096779A1
 ; GENERAL INFORMATION:
 ; APPLICANT: McCall, Catherine A.
 ; APPLICANT: Weber, Eric R.
 ; APPLICANT: Hunter, Shirley Wu
 ; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
 ; TITLE OF INVENTION: AND USES THEREOF
 ; FILE REFERENCE: AL-2-03
 ; CURRENT APPLICATION NUMBER: US/10/218,743
 ; CURRENT FILING DATE: 2002-08-13
 ; PRIOR APPLICATION NUMBER: US/09/292,225
 ; PRIOR FILING DATE: 1999-04-15

; PRIOR FILING DATE: 1998-04-17
 ; PRIOR APPLICATION NUMBER: 09/062,013
 ; PRIOR FILING DATE: 1998-04-17
 ; NUMBER OF SEQ ID NOS: 49
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 34
 ; LENGTH: 1621
 ; TYPE: DNA
 ; ORGANISM: Dermatophagoides farinae
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (14)..(1540)
 ; US-10-218-743-34

Alignment Scores:
 Pred. No.: 8,83e-74 Length: 1621
 Score: 762.00 Matches: 162
 Percent Similarity: 52.36% Conservative: 93
 Best Local Similarity: 33.26% Mismatches: 172
 Query Match: 31.78% Indels: 60
 DB: 9 Gaps: 14

US-10-004-219b-10 (1-445) x US-10-218-743-34 (1-1621)

QY 2 LysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaargPheLeu 21
 Db 113 CGAATCGTATGTTATGTCGAACATGTCGCGTTTATCATAAAGTTGCATCA---TACACA 169
 QY 22 ProLysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMetThr 41
 Db 170 ATTGAAGATATTGATCTTTCAAATGCTACTCATTTGATGTTGTTTCTGCTAAATTCGAT 229
 QY 42 AsnHisGlnLeuSerThr-----ThrGluTrp 50
 Db 230 GAATACAAATACACCATTCACGTTTGTGATTCATTTCAAGATGATCAACATTAACATCATGG 289
 QY 51 AsnAspGluThrLeuTyrGlnGluPheAsnGlyLeuLysMetAsnProLysLeuLys 70
 Db 290 ---GAAACACCGGTATGAACGTTTCAACAACTTGAGATTGAAGATTCAGAAATTCGAC 346
 QY 71 ThrLeuLeuAlaIleGlyGlyTrpAsnGlyThrGlnLysPheThrAspMetValAla 90
 Db 347 ACCATGATTTCATTTGGTGGTGGTATGAAGGTTTCAGAAATATTTCGGATATGGCAGCC 406
 QY 91 ThrAlaAsnAsnArgGlnThrPheValAsnSerAlaIleArgPheLeuArgLysTyrSer 110
 Db 407 AATCCACATATCGTCAGCAATTTGTCATCACTTTTGGACTTTTTCAGAGATACAA 466
 QY 111 PheAspGlyLeuAspLeuAspTrpGluTyrProGlySerGln---GlySerProAlaVal 129
 Db 467 TTCGATGGCTAGATTTCGATTGGGAATATCTGGATCACGGTTAGGCAATCTTAAATC 526
 QY 130 AspLysGluArgPheThrThrLeuValGlnAspLeuAlaAsnAlaPheGlnGluAla 149
 Db 527 GATAACAAAACTATTATTAACATTAGTTAGAGAACTTAAAGAGCAVTTGAACCTTTCGCG 586
 QY 150 GlnThrSerGlyLysGluArgLeuLeuSerAlaAlaValProAlaGlyGlnThrTyr 169
 Db 587 -----TACTTGTGACTGCGCGAGTATCACCCCGTAAAGATAA 625
 QY 170 ValAspAlaGlyTyrGluValAspLysIleAlaGlnAsnLeuAspPheValAsnLeuMet 189
 Db 626 ATTGAGCTAGCTATGAGCTCAAGAATTGAACCAATTTGTCGATGGATGAATGTCATG 685
 QY 190 AlaTyrAspPheHisGlySerTrpGluLysValThrGlyHisAsnSerProLeuTyrLys 209
 Db 686 ACTTATGATTACCATGGCGGATGGGAAATGTTTCGGCCATAATGCTCCGTGTGATAAA 745
 QY 210 ArgGlnGluSerGlyAlaAlaAlaSer---LeuAsnValAspAlaAlaValGlnGln 228
 Db 746 CGACCCGATGAACGGATGAATGTCACATTACTTCAATGTCACTACACCATGCACTAT 805
 QY 229 TrpLeuGlnLysGlyThrProLysSerLysLeuIleLeuGlyMetProThrTyrGlyArg 248

Db 806 TATTTGAACAATGGCGCTACTCGACACAACTGTGTATGGGTTCATTTCTATGGTCT 865
 QY 249 SerPheThrLeuAlaSerSerSerAspThrArgValGlyAlaProAlaThrGlySerGly 268
 Db 866 GCTTGAGCATCGAAGATCGAAGCAAACTGCAACTGGGATCCGCGCAAGGCATGCT 925
 QY 269 ThrProGlyProPheThrLysGluGlyMetLeuAlaTyrTyrGluValCysSerTrp 288
 Db 926 CCTCTGTTTATTACTGGTGAAGAAGGTGTTCTCTCATACATCGAATTTGTGAGTTA 985
 QY 289 LysGlyAlaThrLysGlnArgIleGln---AspGln-----LysValProTyrTrilePhe 305
 Db 986 TTCCAGAAAGAAGATGGCATATTCAATACGATGAATATTACAATGCTCCATACCGATAT 1045
 QY 306 ArgAspAsnGlnTrpValGlyPheAspAspValGluSerPheLysThrLysValSerTrp 325
 Db 1046 AATGATAAAATCTGGTGGTTACGATGATCTGGGTAGTATATCATGCAAGTGGCCITT 1105
 QY 326 LeuLysGlnLysGlyLeuGlyAlaMetValTrpAlaLeuAspLeuAspPheAla 345
 Db 1106 CTCNAAGAATTGGGCGTCTCTGGCGTTATGATATGTCATTGGAAACGATGATTTCAA 1165
 QY 346 GlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeuArgGlnGluLeuSer 365
 Db 1166 GGT---CATTCGGGACCG---AAATATCCATTGTTTGAACAAAGTTTCACAATATGATCAAT 1219
 QY 366 -----LeuPro 367
 Db 1220 GGTGATGAAGAAGAACTCTTACGATGTCTTTGGGCCCCAAGTACAAACACCAACACCA 1279
 QY 368 TyrLeuProSerGlyThrProGluLeuValProLysProGlyGlnProSerGluPro 387
 Db 1280 ACCACCCGCTCACTACT-----TCGACTACCACCAACCGCTACCACCCGAT 1330
 QY 388 GluHisGlyProSerProGlyGlnAspThrPheCysGlnGlyLys----- 402
 Db 1331 AGCACAGCGAAACACCAACAAATACACTAGTATATTGATGGACATTTGATTAATGCTAT 1390
 QY 403 AlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyrSerCys----- 419
 Db 1391 AAACAAGGTATTCTTCCACATCCAACTGATGTCTATAAATATTAGTTTGTGAATATT 1450
 QY 420 -----AlaAlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPhe 435
 Db 1451 GCCACACCAACCGGTGGTGGTGTCACATATTGATTTGTCACAAAGGAAGTATAGTGG 1510
 QY 436 SerAsnSerCysLysCysCys 442
 Db 1511 CACGCAACATTAAATAATGT 1531

RESULT 13

US-10-218-743-36/c
 ; Sequence 36, Application US/10218743
 ; Publication No. US20030096779A1
 ; GENERAL INFORMATION:
 ; APPLICANT: McCall, Catherine A.
 ; APPLICANT: Hunter, Shirley Wu
 ; APPLICANT: Weber, Eric R.
 ; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
 ; FILE OF INVENTION: AND USES THEREOF
 ; FILE REFERENCE: AL-2-C3
 ; CURRENT APPLICATION NUMBER: US/10/218,743
 ; CURRENT FILING DATE: 2002-08-13
 ; PRIOR APPLICATION NUMBER: US/09/292,225
 ; PRIOR FILING DATE: 1999-04-15
 ; PRIOR APPLICATION NUMBER: 60/098,909
 ; PRIOR FILING DATE: 1998-09-02
 ; PRIOR APPLICATION NUMBER: 60/085,295
 ; PRIOR FILING DATE: 1998-05-13
 ; PRIOR APPLICATION NUMBER: 60/098,565
 ; PRIOR FILING DATE: 1998-04-17
 ; PRIOR APPLICATION NUMBER: 09/062,013

Alignment Scores:		
Pred. No.:	3.76e-71	Length: 1608
Score:	738.00	Matches: 156
Percent Similarity:	54.71%	Conservative: 88
Best Local Similarity:	34.98%	Mismatches: 143
Query Match:	30.78%	Indels: 60

Db	573	CTGAAGAAATAGGCGTTCTTGGTGTATGGTTGGTCATTGGAAAAATGATGATTTCAA	514
		::: ::: ::: ::: ::: ::: ::: :::	
Qy	346	Gly	346
Db	513	GGTCACTGGGACCGAAAAATCCATTGTTGAACAAAGTTTCATAATATGATTATATGGCGAT	454
Qy	347	-----PheSerCys-----	353
Db	453	GAAGAAAGAACTCTTTCCGAATGCATTTTGGGTCCAAAGTACAACGACACCAACTCCAACGCAGC	394
Qy	353	gTyxProLeuIleGlnThrLeuArgGlnGluLeuSerLeuProTyrLeuProSerGlyTh	373
		::: ::: ::: ::: ::: ::: ::: :::	
Db	393	ACACCCACAAACCCCGACTACACGCCACCAACTCTTCTCCACAC-ACCCECGACAACAAC	335
Qy	373	rProGluLeuGluValPro---LysProGlyGlnProSerGluProGluHisGlyProSe	392
		::: ::: ::: ::: ::: ::: ::: :::	
Db	334	CCCTCTCCACCAACCCCGACCAACAACCCCTTCTCCACCAACCCGACCAACAACACTCCTTC	275
Qy	392	rProGlyGlnAspThr	397
Db	274	TCCACCAACCAACA	259

Search completed: June 30, 2003, 03:58:16
Job time : 186.267 secs

THIS PAGE BLANK (USPTO)

REFERENCE

1 Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636

REFERENCE

2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20493374
PUBMED 11042159

REFERENCE

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuami, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamaoka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadoro, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischnann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hoffmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schombach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whitaker, C., Wilming, L., Wyehaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409 (6821), 685-690 (2001)
MEDLINE 21085660
PUBMED 11217851

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

5 (bases 1 to 1581)
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Hori, F., Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ono, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Schriml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamanaka, I., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
TITLE Direct Submission
JOURNAL Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gscc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222.

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

COMMENT
Fax: 81-45-503-9216)
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'-GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTT 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5'-GAGAGAGATTCGAGTTAAATTAATTAATCCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.

Host: SOLR

Location/Qualifiers

1. .1581
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="PANTOM DB:2300002L19"
/db_xref="MGD:MGJ:1904571"
/db_xref="taxon:10090"
/clone="2300002L19"
/sex="male"
/tissue type="tongue"
/clone lib="RIKEN full-length enriched mouse cDNA library"
/dev stage="adult"
230. .1420
/note="data source:SPTR, source key:Q9H3V8, evidence:ISS homolog to CHITOTRIOSIDASE PRECURSOR putative"

CDS

/codon start=1
/protein id="BAB26025.1"
/db_xref="GI:12843548"
/translation="MVQSLAWGVMTLLMVQMSAAKLVCLYTNNSQVTEAVRFPFR DVNPLCTHVIAPAFQMDNHLSTVEHNDLLYLQELNSLKTGNPKLKTLLAVGGWTFG TOKFTDMVATSNROTFFVKSPDTQFGLDLDWEPGPGSGPTVDKERTALIQ DLAKAFQEAQSSGKRLTAAVSPDRGLVDAGVEVDKIAQSLDFINLMADHSSLL EKTGHNPLYKQGESGAARQNDVAATLWLGKTPASKLLGMPYGRSFTLASS SUNGVGATKPGAGPYTKQGVLYAYEACSWKRRHREDOKVTFAPDNQWVSFDD VESFKAALMGSTPTLETSLPTTTVEBGGSSRAVLQAWCLEPLANVVPGAKFLEPH PTFP"

polyA_signal

1559. .1564

polyA_site

1581
/note="putative"
BASE COUNT 360 a 454 c 423 g 344 t
ORIGIN

Alignment Scores:

Pred. No.: 3,148-134 Length: 1581
Score: 1471.00 Matches: 281
Percent Similarity: 70.11% Conservative: 31
Best Local Similarity: 63.15% Mismatches: 53
Query Match: 61.34% Indels: 81
DB: 11 Gaps: 2

US-10-004-219B-10 (1-445) x AK009012 (1-1581)

QY 1 AlaLysLeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPhe 20

293 GCAAACTGGTCTGCTACTCTACCAACCTGGTCCAGTACCGACGAGGACGTCGTTTC 352

Db 21 LeuProLysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMet 40

353 TTTCAGGAGATGGATGATCCCAACCTGTGTATCCACGTCATCTTTGCTTCTGGAATG 412

QY 41 ThrAsnHisGlnLeuSerThrThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsn 60

413 GACAACCATCAGCTCAGACCTGGAGACATGACAACTCTCTTACGAGGAGCTGAAC 472

QY 61 GlyLeuLysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPhe 80

473 AGCCTAAAGACTAAGAACCCCAAGCTCAAGACCTGTGTAGCCGTTGGAGCTGGACCTTT 532
QY
81 GlyThrGlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsn 100
DB
533 GGTACCCAGAAAGTTCACAGACATGGTGGCCACCGCCAGCAACCGCAGACCTTTGTGAAG 592
QY
101 SerAlaIleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTyrGluTyr 120
DB
593 TCAGGCCCTAAGTTTCTGGCGACTCAAGGTTTGTATGGCTTGACCTTGACCTGGAGTTC 652
QY
121 ProGlySerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAsp 140
DB
653 CCAGGTGGACGTGGGAGCCCAACAGTAGACAAGAGAGATTACACGCCCTGATACAGGAC 712
QY
141 LeuAlaAsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuSer 160
DB
713 TTGGCCAAAGCCTTCAGGAGGAAGCCCAAGTCTCAGGGAAGGAAGCCCTCTCTTGACT 772
QY
161 AlaAlaValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAla 180
DB
773 GCAGCTGTACCAAGTATCGAGCGCTGGTGGATGTGGCTAGAGGTGGACAGATTGCC 832
QY
181 GlnAsnLeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysVal 200
DB
833 CAGAGCTTGGAATTCATCAACCTTATGGCTTACGACTTCACAGCTCTCTTGGAAGAC 892
QY
201 ThrGlyHisAsnSerProLeuTyrLysArgGlnGluSerGlyAlaAlaAlaSerLeu 220
DB
893 ACAGGGCATAAACAGCCCTCTACAAAAGCAAGGAGAAAGTGGGACACCCCTGAGCAA 952
QY
221 AsnValAspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuIle 240
DB
953 AACGTGGATGCTGTGTGAGCTCTGGCTGCAAGAGGGACCCAGCCAGCAAACTGATC 1012
QY
241 LeuGlyMetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgVal 260
DB
1013 CTGGCATGCTACCTATGAGCGCTCTTTCACTTGGCTCTCTCGTCAGACAATGGAGTT 1072
QY
261 GlyAlaProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyMetLeu 280
DB
1073 GGGGCCCCAGCACAGGGCTGTGGTCCCGAGGCCCTATACGAAGCAAAAGGGGCTCTG 1132
QY
281 AlaTyrTyrGluValCysSerTrpLysGlyAlaThrLysGlnArgIleGlnAspGlnLys 300
DB
1133 GCTTACTATGAGCGCTCTCTGGAG-----GAAAGACACAGATCCAGGACCAAG 1186
QY
301 ValProTyrIlePheArgAspAsnGlnTrpValGlyPheAspAspValGluSerPheLys 320
DB
1187 GTGCCTTACGCTTCCAGGACCAACCACTGGTGGTGGCTTTGACGACGTGGAAAGCTTCAAA 1246
QY
321 ThrLysValSerTyrLeuLysGlnLysGlyLysGlyAlaMetValTrpAlaLeuAsp 340
DB
1247 GC----- 1248
QY
341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeu 360
DB
1248 ----- 1248
QY
361 ArgGlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuValProLys 380
DB
1248 ----- 1248
QY
381 ProGlyGlnProSerGluProGluHisGlyProSerProGlyGlnAspThrPheCysGln 400
DB
1249 -----CAA 1251
QY
401 GlyLysAlaAspGlyLeuTyrProAsnProArgGluArgSerSerPheTyrSerCysAla 420
DB
1252 GGCAAGCTGATGGGTCTACCCCAACCTCGAGAGCAGTCCACTTACTACAACTGTGGA 1311
QY
421 AlaGlyArgLeuPheGlnGlnSerCysProThrGlyLeuValPheSerAsnSerCysLys 440

DB 1312 GGAGGCGCGCTGTTCCAGCAGAGCTGTCTCCAGCGCTGTGTGTTTAGAGCCCTCTTGC AAA 1371
QY 441 CysCysThrTrpAsn 445
DB 1372 TGTGTACTCTGAGC 1386
RESULT 2
LOCUS AK008633 1533 bp mRNA linear HTC 19-JAN-2002
DEFINITION Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:2200003E03:chitinaae, acidic, full insert sequence.
ACCESSION AK008633
VERSION AK008633.1 GI:12842941
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (strain: C57BL/6J) adult male stomach cDNA to mRNA, clone lib:RIKEN full-length enriched mouse cDNA library
clone:2200003E03.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE 4
AUTHORS Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Giesi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bulc, C., Hofmann, M., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Fletcher, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombarts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohseki, S. and Hayashizaki, Y.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409 (6821), 685-690 (2001)
MEDLINE 21085660
PUBMED 11217851
REFERENCE 5
AUTHORS Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,


```

putative"
/codon_start=1
/protein_id="BAB25878.1"
/db_xref="GI:12843149"
/db_xref="MGI:1932052"
/translation="MQNEITTEWMDVLYKAFNDLKRNNSKLTLAIGWNNFOTA
PFTTMYSTQNRQTFITTSVIFKRLQYQFDGLDWEYPGSRGSPQDKHLFTVLVKEM
REAFEOEARIESNPRMLMTAAVAGGISNIOAGYEIPELSKLDFIHVMYVDLHGSWEG
YTGENSEPLKYPTPETGSLVNDVYVMYKNGAPAEKLIUVGFPEYGHAFILRNPSD
NLGATSGDGPAGPYTRQAGFWAYYEICTFLRSGATEVNDASQ"
polyA_signal
1520..1525
/notes="putative"
1535
polyA_site
/notes="putative"
1535
BASE COUNT 379 a 381 c 373 g 402 t
ORIGIN

Alignment Scores:
Pred. No.: 8,07e-118 Length: 1535
Score: 1304.50 Matches: 236
Percent Similarity: 67.11% Conservative: 70
Best Local Similarity: 51.75% Mismatches: 129
Query Match: 54.40% Indels: 21
DB: 11 Gaps: 3

US-10-004-219B-10 (1-445) x AK008757 (1-1535)
QY 3 LeuValCysThrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPheLeuPro 22
DB 86 CTGATATGCTATTTCACCAACTGGGCCCAAGTATCGGCCAGGCTGGGGAGCTTCAAGCCT 145
QY 23 LysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMetThrAsn 42
DB 146 GATGACATTAACCCCTGCTGTGTACTACCTGATCTATGCCCTTGTCTGGGATGCGAAC 205
QY 43 HisGlnLeuSerThrThrGlnTrpAsnAspGluThrLeuTyrGlnGluPheAsnGlyLeu 62
DB 206 AATGAGATCACCACCATAGAAATGGAATGATGTTACTCTCTATAAAGCTTTCATGACTTG 265
QY 63 LysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyTyrPheAsnGlyThr 82
DB 266 AAAACAGGACGACGACCAATGAAACCCCTCGCAATGGAGCTGGAACTTTGGAATC 325
QY 83 GlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsnSerAla 102
DB 326 GCTCCTTTTCACTACCATGTTTCCACTTCTCAGAACCGCCAGACCTTCATTACCTCAGTC 385
QY 103 IleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyrProGly 122
DB 386 ATCAAAATTTCTGGCTCAGTATGGTGTGTGATGGACTGGACTGGGAAATACCCAGGC 445
QY 123 SerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAspLeuAla 142
DB 446 TCACGTGGGAGGCCCTCCTCAGGACAAAGCATCTCTTCACTGCTCCTGGTGAAGGAATGCGT 505
QY 143 AsnAlaPheGlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuSerAlaAla 162
DB 506 GAAGCTTTTGAGCAGGAGGCTATTGAGACGACAGCCAGCTGATGCTGTACTGCTGT 565
QY 163 ValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAlaGlnAsn 182
DB 566 GTAGCTGGTGGGATTTCCAAATCCAGGCTGGCTGATGAGATCCCTGAACCTTCTTAAGTAC 625
QY 183 LeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysValThrGly 202
DB 626 CTGGATTTTCATCATGCTAAGCATATGATACCTCCATGGCTCTCGGAGGGCTACATGGG 685
QY 203 HisAsnSerProLeuTyrLysArgGlnGluGluSerGlyValAlaAlaSerLeuAsnVal 222
DB 686 GAGATAGTCTCTTTACAAATACCCTACTGAGACTGGTAGCAATGCCCTACCTCAATGTG 745
QY 223 AspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuIleLeuGly 242

```

```

DB 746 GATTATGTCACTGAACCTATTGGGAAGAACCAATGGAGCCCGCAGCTGAGAAGCTCATTTGTGGA 805
QY 243 MetProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgValGlyAla 262
DB 806 TTCCAGAGTAGTGGACACGCCCTTCATCTCGAAGAACCCCTCTGTAATAGGAATTTGTGTC 865
QY 263 ProAlaThrGlySerGlyThrProGlyProPheThrLysGluGlyMetLeuAlaTyr 282
DB 866 CTACTCTGTGTGGTGGCCCTCTGGGCCCTATACACGACAGCTGGTTCCTGGGCTTAC 925
QY 283 TyrGluValCysSerTrp-----LysGlyAlaThrLysGlnArgIleGlnAspGlnLys 300
DB 926 TATGAGATTTCACCTTTCTGAGAAAGTGAGGACCTAGAGTCTGGGATGCCCTCCCAATAA 985
QY 301 ValProTyrIlePheArgAspAsnGlnTrpValGlyPheAspValGluSerPheLys 320
DB 986 GTGCCCTATGCTTAAAGGCCAACGAGTGGCTTGGCTATGACCAATATCAAGAGCTTCAGT 1045
QY 321 ThrLysValSerTyrLeuLysGlnLysGlyLeuGlyGlyAlaMetValTrpAlaLeuAsp 340
DB 1046 GTTAAGGCTCAGTGGCTTAAGCAGAACAAATTTTGGAGGTGCCATGATCTGGGCCATTGAC 1105
QY 341 LeuAspAspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeu 360
DB 1106 CTTGATGACTTCACTGGCTCTTTCTGTGATCAGGGAATAATTTCTCTGACTTCTACTTTG 1165
QY 361 ArgGlnGluLeuSerLeuProTyrLeuProSerGlyThrProGluLeuGluValProLys 380
DB 1166 AACAAAGCCCTGGCATATPCCATGAAGTTGCCAGCTCTGACGTG-----Glu 1213
QY 381 ProGlyGlnProSerGluPro-----CCTTCGAGCCAGTACTCTCCTCAGGAAGTGGGAGTGGAGC 1264
DB 1214 -----CCTTCGAGCCAGTACTCTCCTCAGGAAGTGGGAGTGGGAGTGGAGC 1264
QY 389 HisGlyProSerProGlyGlnAspThrPheCysGlnGlyLysAlaAspGlyLeuTyrPro 408
DB 1265 TCCGGAGGAAGCTCTGGAGCGAGTGGATTCTGTGCCGACAAAGCAGATGGCCTTACCT 1324
QY 409 AsnProArgGluArgSerPheTyrSerCysAlaAlaGlyArgLeuPheGlnGlnSer 428
DB 1325 GTGGCAGATGACAGAAATGCTTTTGGCAGTGCATCAATGGAAATCATTACACGACGAT 1384
QY 429 CysProThrGlyLeuValPheSerAsnSerCysLysCysCysThrTrp 444
DB 1385 TGTCAAGCAGGCTTGTTTTGTATACCACTGATGTAATTTGTCGCACTGG 1432

RESULT 4
AK008650
LOCUS
DEFINITION
Mus musculus adult male stomach cDNA, RIKEN full-length enriched
library, clone:220008L17:chitinase, acidic, full insert sequence.
ACCESSION
AK008650
VERSION
AK008650.1 GI:12842971
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (strain:C57BL/6J) adult male stomach cDNA to mRNA,
clone lib:RIKEN full-length enriched mouse cDNA library
clone:220008L17.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
PUBMED
10349636
REFERENCE
2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374

```


/clone_lib="NCI CGAP Lu33"
/tissue_type="pooled lung tumors"
/lab_host="DH10B (phage-resistant)"
/note="organ: lung; Vector: pT73D-Pac (Pharmacia) with a
modified polylinker; Site 1: NotI; Site 2: EcoRI; 1st
strand cDNA was prepared from mRNA obtained from pooled
lung tumors with a Not I - oligo(dT) primer [5',
TGTTACCAATCTGAAGTGGAGCGCCCTCTGTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not
I and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 202 a 248 c 230 g 157 t

ORIGIN

Alignment Scores:
Pred. No.: 2,76e-84 Length: 837
Score: 960.50 Matches: 185
Percent Similarity: 79.38% Conservative: 19
Best Local Similarity: 71.98% Mismatches: 48
Query Match: 40.05% Indels: 5
DB: 13 Gaps: 2

US-10-004-219B-10 (1-445) x B1412518 (1-837)

QY 191 TyrAspPheHisGlySerTrpGluLysValThr-GlyHisAsnSerProLeuTyrLysAr 210
DB 5 TACGACTTCCAGCTCTCTGGAAAG---ACACAGGCATACAGCCCTCTACAAAG 61
QY 210 ggInGluSerGlyAlaAlaSerLeuAsnValAspAlaAlaValGlnGlnTrpLe 230
DB 62 GCAAGAGAAAGTGGGCGAGCCGCTGAGCAAAACGTGGATGCTGTGACGCTCTGGCT 121
QY 230 uGlnLysGlyThrProAlaSerLysLeuLeuLeuGlyMetProThrTyrGlyArgSerPh 250
DB 122 GCAGAGGAGGAGCCGCCAGCAGCAAACTGATCTCTGGCATGCTTACATATGAGCGCTCTTT 181
QY 250 eThrLeuAlaSerSerAspThrArgValGlyAlaProAlaThrGlySerGlyThrPr 270
DB 182 CACCTTGGCTCTCTGTCAGCAATGAGTGGGCCCCCAGCCAGGCGCTGGTGGCCCC 241
QY 270 oGlyProPheThrLysGluGlyGlyMetLeuAlaTyrTyrGluValCysSerTrpLysGl 290
DB 242 AGGCCCTATACGAAGGAGGAGGCTCTGCTTACTATGAGGCTGCTCTCGAAG-- 299
QY 290 yAlaThrLysGlnArgIleGlnAspGlnLysValProTyrIlePheArgAspAsnGlnTr 310
DB 300 ----GAAAGACACAGAAATCGAGGAGGAGGAGGCTTACGCGCTTCCAGGCAACCCAGTG 355
QY 310 pValGlyPheAspAspValGluSerPheLysThrLysValSerTyrLeuLysGlnLysGl 330
DB 356 GGTGAGCTTTGACGAGCTGGAAGGCTTCAAGCCAGGCTGCTTCTTCAAGCAAGAGG 415
QY 330 yLeuGlyGlyAlaMetValTrpAlaLeuAspLeuAspPheAlaGlyPheSerCysAs 350
DB 416 GCTGGGAGGAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 475
QY 350 nGlnGlyArgTyrProLeuIleGlnThrLeuArgGlnGlnLeuSerLeuProTyrLeuPr 370
DB 476 CCAGGCGCCGCTACCTCTCATCCGACACTACGCGAGGAACTAAATCTTCCATCCGAGAC 535
QY 370 oSerGlyThrProGluLeuGluValProLysProGlyGlnProSerGluProGluHigl 390
DB 536 TCCAGGAGGCCAGAACAGATATACCTGAGCCAGCCGCCCATCTTATGCCAGAGCAGGG 595
QY 390 yProSerProGlyGlnAspThrPheCysGlnGlyLysAlaAspGlyLeuTyrProAsnPr 410
DB 596 ACCAGCCAGGCGTAGATAACTTCTGCCAGGCAAGCTGATGGGTCTACCCCAACC 655
QY 410 oArgGluArgSerSerPheTyrSerCysAlaAlaGlyArgLeuPheGlnGlnSerCysPr 430
DB 656 TGGAGACGAGTCCACTTACTACAACTGTGGAGGAGGGCGCTGTTCCAGCAGAGCTGTAC 715

QY 430 oThrGlyLeuVal-PheSerAsnSerCysLysCysCysThrTrpAsn 445
DB 716 TCCAGGCTGTGGTTAGAGCCCTCGCAATGTGTACCTGGAGC 762

RESULT 8

BQ231098

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 892)

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM13305 row: k column: 05

High quality sequence stop: 772.

Location/Qualifiers

1..892

/organism="Mus musculus"

/db_xref="taxon:10090"

/clone="IMAGE:6051580"

/clone_lib="NCI CGAP St1"

/lab_host="DH10B (T1-resistant)"

/note="Organ: stomach; Vector: pCMV-SPORT6; Site 1: SalI;

Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.

Average insert size 1.77 kb. Library constructed by Life

Technologies. Note: this is a NCI CGAP Library."

BASE COUNT 219 a 226 c 216 g 230 t

ORIGIN

Alignment Scores:

Pred. No.: 7.28e-79 Length: 892

Score: 906.00 Matches: 161

Percent Similarity: 74.18% Conservative: 43

Best Local Similarity: 58.58% Mismatches: 71

Query Match: 37.78% Indels: 0

DB: 14 Gaps: 0

US-10-004-219B-10 (1-445) x BQ231098 (1-892)

QY 3 LeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPheLeuPro 22

DB 42 CTGATATGCTATTTCCCAACTGGGCCCCAGTATCGCCAGGCTTCAGGCCT 101

QY 23 LysAspLeuAspProSerLeuCysThrHisLeuIleTyrAlaPheAlaGlyMetThrAsn 42

DB 102 GATGACATTAACCCCTGCTGTACTACCTGATCTATGCTTTGCTGGATGAGAAC 161

QY 43 HisGlnLeuSerThrGluTrpAsnAspGluThrLeuTyrGlnGluPheAsnGlyLeu 62

DB 162 AATGAGATCACCAACCAAGATGATGTTACTCTCTATAAAGCTTCAATGACTTG 221

QY 63 LysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnGlyThr 82

DB 222 AAAAAACAGAAACAGAACTGAAACCCCTCTGGCAATTGGAGCTTGAACCT 281

QY 83 GlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsnSerAla 102

```
Db 282 GCTCCTTTTCACTACCATGTTTCCACTTCTCAGAACCGCAGACCTTCATTACCTCAGTC 341
Qy 103 IleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyrProGly 122
Db 342 ATCAAAATTTCTGCGTCAGTATGGTGTGATGGAGCTGGACCTGGAGCTGGGAATACCCAGGC 401
Qy 123 SerGlnGlySerProAlaValAspLysGluAtpGpHeThrThrLeuValGlnAspLeuAla 142
Db 402 TCACGTGGAGAGCCCTCTCAGCAAGCATCTCTTCACTGTCTGTGTGAGGAATGCCT 461
Qy 143 AsnAlaPheGlnGlnGlnAlaGlnThrSerGlyLysGluArgLeuLeuLeuSerAlaAla 162
Db 462 GAAGCTTTTGGAGCAGGAGCTATTGAGAGCAACAGCCAGACCTGATGTTACTGTCTGCT 521
Qy 163 ValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysLysLeuAlaGlnAsn 182
Db 522 GTAGCTGGTGGGATTTCCAAACATCCAGGCTGGCTATGAGATCCCTGAACTTTCTAAGTAC 581
Qy 183 LeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysValThrGly 202
Db 582 CTGGATTTTCACTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 641
Qy 203 HisAsnSerProLeuTyrLysArgGlnGlnGluSerGlyAlaAlaAlaSerLeuAsnVal 222
Db 642 GAGATAGTCTCTTTACAAATACCTCTAGACTGGTAGCAATGCCTACCTCAATGTG 701
Qy 223 AspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuLeuGly 242
Db 702 GATTATGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 761
Qy 243 MetProThrTyrGlyArgSerPheThrLeuAlaSerSerAspThrArgValGlyAla 262
Db 762 TTCCAGAGTATGGACACACTTCTCTGAGAAACCCCTCTGATGATGATGATGATGATGATGAT 821
Qy 263 ProAlaThrGlySerGlyThrProGlyProPheThrLysGluGly 277
Db 822 CCTACCTCTGGTGTGCTGGCCCTGCTGGCCCTATACCCAGACAGGCG 866

RESULT 9
BG867815 838 bp mRNA linear EST 29-MAY-2001
LOCUS 602786336F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4912544 5'
DEFINITION mRNA sequence.
ACCESSION BG867815
VERSION BG867815.1 GI:14218355
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 838)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: L14M10816 row: 9 column: 09
High quality sequence stop: 824.
Location/Qualifiers
1 .838
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4912544"
/clone_lib="NCI_CGAP_SG2"
```

FEATURES
source

```
/lab host="DH10B (T1 phage-resistant)"
/Note="Organ: salivary gland; Vector: pCMV-SPORT6; Site 1:
Note1: Site 2: SalI; Cloned unidirectionally. Primer: Oligo
dT. Average insert size 1.3 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 209 a 211 c 198 g 220 t
ORIGIN
Alignment Scores: 2,876-74 Length: 838
Pred. No.: 859.00 Matches: 155
Score: 73.31% Conservative: 40
Best Local Similarity: 58.27% Mismatches: 71
Query Match: 35.82% Indels: 0
DB: 12 Gaps: 0
US-10-004-219B-10 (1-445) x BG867815 (1-838)
Qy 3 LeuValCysTyrPheThrAsnTrpAlaGlnTyrArgGlnGlyGluAlaArgPheLeuPro 22
Db 40 CTGATATGCTATTTTCCCAACTGGGCCAGTATGGCCAGGTCCTGGGAGCTTCAAGCCT 99
Qy 23 LysAspLeuAspProSerLeuCysThrHisLeuLysTyrAlaPheAlaGlyMetThrAsn 42
Db 100 GATGACATTAACCCCTGCTGTGTACTCACCTGATCTATGCTGCTTTGCTGGGATGCAGAAC 159
Qy 43 HisGlnLeuSerThrGlnTrpAsnAspGluThrLeuTyrGlnGluPheAsnGlyLeu 62
Db 160 AATGAGATCACCACCATGAGATGGAATGATGATGATGATGATGATGATGATGATGATGATG 219
Qy 63 LysLysMetAsnProLysLeuLysThrLeuLeuAlaLysGlyTyrTrpAsnPheGlyThr 82
Db 220 AAAACACAGGACAGCAAACTGAAACCTCTCTGGCAATTGGAGGCTGGAACTTTGGAACT 279
Qy 83 GlnLysPheThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsnSerAla 102
Db 280 GCTCCTTTTCACTACCATGCTTTTCCACTTCTCAGAACCCGACAGCTTCAATACCTCAGTC 339
Qy 103 IleArgPheLeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrpGluTyrProGly 122
Db 340 ATCAAAATTTCTGCGTCAGTATGGTGTGATGGAGCTGGACCTGGAGCTGGGAATACCCAGGC 399
Qy 123 SerGlnGlySerProAlaValAspLysGluArgPheThrThrLeuValGlnAspLeuAla 142
Db 400 TCACGTGGAGGACCTCTCAGCAAGCATCTCTTCACTGTCTGTGTGAGGAATGCCT 459
Qy 143 AsnAlaPheGlnGlnGlnAlaGlnThrSerGlyLysGluArgLeuLeuLeuSerAlaAla 162
Db 460 GAACTTTTGGACGAGGAGCTATTGAGAGCAACAGCCAGACCTGATGCTGTCTGCTGTCT 519
Qy 163 ValProAlaGlyGlnThrTyrValAspAlaGlyTyrGluValAspLysLysLeuAlaGlnAsn 182
Db 520 GTAGCTGGTGGGATTTCCAACTCCAGGCTGGCTGCTATGAGATCCCTGAACTTTCTAAGTAC 579
Qy 183 LeuAspPheValAsnLeuMetAlaTyrAspPheHisGlySerTrpGluLysValThrGly 202
Db 580 CTGGATTTTCACTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 639
Qy 203 HisAsnSerProLeuTyrLysArgGlnGlnGluSerGlyAlaAlaAlaSerLeuAsnVal 222
Db 640 GAGATAGTCTCTTTTACAAATACCTCTAGACTGGTAGCAATGCCTTACCTCAATGTG 699
Qy 223 AspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuLeuGly 242
Db 700 GATTATGTCATGAACTATTGGAGAACATGGAGCCCGCCAGCTGAGAGCTCATTTGTTGA 759
Qy 243 MetProThrTyrGlyArgSerPheThrLeuAlaSerSerAspThrArgValGlyAla 262
Db 760 TTCCAGAGATTTGGACACACCTTCTCTGAGAAACCCCTCTGATAATGGAATTTGGTGCC 819
Qy 263 ProAlaThrGlySerGly 268
Db 820 CCTACTCTGCTGGTGAATGGC 837
```


Tissue Procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM10840 row: 1 column: 19
 High quality sequence stop: 845.

FEATURES

source
 Location/Qualifiers
 1..880
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:4921890"
 /clone_lib="NCI CGAP S2"
 /lab_host="DH10B (T1 phage-resistant)"
 /note="Organ: salivary gland; Vector: pCMV-SPORT6; Site: 1:
 NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo
 dt. Average insert size 1.3 kb. Constructed by Life
 Technologies. Note: this is a NCI CGAP Library."
 BASE COUNT 218 a 220 c 211 g 231 t
 ORIGIN

Alignment Scores:
 Pred. No.: 3,34e-69 Length: 880
 Score: 808.00 Matches: 151
 Percent Similarity: 71.22% Conservative: 42
 Best Local Similarity: 55.72% Mismatches: 78
 Query Match: 33.69% Indels: 1
 DB: 12 Gaps: 0

US-10-004-219B-10 (1-445) x BG871468 (1-880)

QY 3 LeuValCysThrPheThrAsnTrpAlaGlnTyArgGlnGlyGluAlaArgPheLeuPro 22
 Db 67 CTGATATGTTATTCACCACTGGGCCCAAGTATCGGCCAGCTTCAAGCCT 126
 QY 23 LysAspLeuAspProSerLeuCysThrHisLeuIleTyAlaPheAlaGlyMetThrAsn 42
 Db 127 GATGACATTAACCCCTGCTGTGTACTACCTGATCTATGACCTTTGCTGGGATGCGAAC 186
 QY 43 HisGlnLeuSerThrThrGluTrpAsnAspGluThrLeuTyGlnGluPheAsnGlyLeu 62
 Db 187 AATGAGATCACCACCATAGATGGAATGATGTTACTCTCTATAAGCTTTCATGACTTG 246
 QY 63 LysLysMetAsnProLysLeuLysThrLeuLeuAlaIleGlyThrAsnPheGlyThr 82
 Db 247 AAAAAAGGAGAACAGCAAACTGAAACCCCTCTGGCAATTTGGAGGCTGGAACTTTGGA 306
 QY 83 GlnLysPheThrAspMetValAlaThrAlaAsnArgGlnThrPheValAsnSerAla 102
 Db 307 GCTCCTTTCACTACCATGTTTCCATCTTCAGAACCGCCAGACCTTCATACCTCAGTC 366
 QY 103 IleArgPheLeuArgLysTySerPheAspGlyLeuAspLeuAspTrpGluTyProGly 122
 Db 367 ATCAAAATTTCTGCTCAGTATGCTTTGATGGACTGACCTGACCTGGGAATACCCAGGC 426
 QY 123 SerGlnGlySerProAlaValAspLysGluArgPheThrLeuValGlnAspLeuAla 142
 Db 427 TCACGTGGGAGCCCTCTCAGGACCAAGCATCTCTTCACTGCTGTGGTGAAGGAATGCGT 486
 QY 143 AsnAlaPheGlnGlnGlnAlaGlnThrSerGlyLysGluArgLeuLeuLeuSerAlaAla 162
 Db 487 GAAGCTTTTGAGCAGGAGGCTATTGAGCAACAGCCAGCTGATGTTGCTGCTGCT 546
 QY 163 ValProAlaGlyGlnThrTyValAspAlaGlyTyGluValAspLysIleAlaGlnAsn 182
 Db 547 GTAGCTGGTGGGATTTCCAACTCCAGGCTGGCTGATGAGATCCCTGAACTTTCTAAGTAC 606
 QY 183 LeuAspPheValAsnLeuMetAlaTyArgPheHisGlySerTrpGlnLysValThrGly 202
 Db 607 CTGGATTTTCACTCATGTCATGACATATGACCTCCATGGCTCTCTGGGAGGGCTACACTGGG 666

QY 203 HisAsnSerProLeuTyLysArgGlnGluSerGlyAlaAlaSerLeuAsnVal 222
 Db 667 GAGATAGTCTCTTTTACAAATACCTACTGAGACTGGTAGCAATGCCTACCTCAATGTG 726
 QY 223 AspAlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuLeuGly 242
 Db 727 GATTATGTCATGACTATTGGGAAGAACCAATGGAGCCAGTGAAGGTCATTGCTGGAT 786
 QY 243 MetProThrTyArgSerPheThrLeuAlaSerSerSerAspThrArgValGlyAla 262
 Db 787 TCAAGAGTATTGGTGAACACCTTCATCTGAGAACCTCTGATCTG-ATTGGTGCC 845
 QY 263 ProAlaThrGlySerGlyThrProGlyProPhe 273
 Db 846 CCTAATCTGGTGGTGGGCTGTGGGCTATTTC 878

RESULT 15

BM458470

LOCUS

DEFINITION

BM458470

VERSION

KEYWORDS

SOURCE

ORGANISM

human

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BM458470

AGENCY

5' mRNA sequence.

BM458470

EST.

human

human sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1098)

NIH-MGC http://mgi.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs@mail.nih.gov

Tissue Procurement: Lou Staudt

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM12126 row: k column: 21

High quality sequence stop: 683.

BASE COUNT 282 a 274 c 276 g 264 t 2 others

ORIGIN

Alignment Scores:

Pred. No.: 7,44e-69 Length: 1098

Score: 806.00 Matches: 162

Percent Similarity: 71.43% Conservative: 58

Best Local Similarity: 52.60% Mismatches: 75

Query Match: 33.61% Indels: 16

DB: 13 Gaps: 5

US-10-004-219B-10 (1-445) x BM458470 (1-1098)

QY 66 AsnProLysLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPheGlyThrGlnLysPhe 85

Db 4 AATCCCAACTGAAATTTCTTTGTC-ATTGAGGGAGCTG-TTTGGTTCCAAAGGTTTC 61

```
QY      86 ThrAspMetValAlaThrAlaAsnAsnArgGlnThrPheValAsnSerAlaIleArgPhe 105
Db      ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
62 CACCCTATGTTGGATTCTTACATCACCGCTTGGAAATTCATTAACTCCATAATCCTGTTT 121
QY      106 LeuArgLysTyrSerPheAspGlyLeuAspLeuAspTrrpGluTyrProGlySerGlnGly 125
Db      ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
122 CTGAGGAACCATAACTTTGATGACTGGATTAAGCTGGATCTACCCAGATCAGAAA--- 178
QY      126 SerProAlaValAspLysGluArgPheThrThrLeuValGlnAspLeuAlaAsnAlaPhe 145
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
179 -----GAAACACTCATTTCACTGTGCTGATTCATGAGTTAGCAGAAGCCTTT 226
QY      146 GlnGlnGluAlaGlnThrSerGlyLysGluArgLeuLeuLeuSerAlaAlaValProAla 165
Db      ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
227 CAGAAGGACTTCACAAANTCCACCAAGGAAGGCTCTCTTGACTGGGGCGTATCTGCA 286
QY      166 GlyGlnThrTyrValAspAlaGlyTyrGluValAspLysIleAlaGlnAsnLeuAspPhe 185
Db      ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
287 GGGAGGCAAAATGATTGATAACAGCTATCAAGTTGAGAAACTGCAAAAGATCTGGATTTC 346
QY      186 ValAsnLeuMetAlaTyrAspPheHisGlySerTrrpGluLys-----ValThrGlyHis 203
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
347 ATCAACCTCCTGTCTCTTCACTTCCATGGGTCTTGGGAAAGCCCTTATCACTGTGCCAC 406
QY      204 AsnSerProLeuTyrLysArgGlnGlnGluSerGlyValAlaAlaSerLeuAsnValAsp 223
Db      ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
407 AACAGCCCTCTGAGCAAGGGGTGGCAGGACAGAGGGCCCAAGCTCTACTACATGTGGAA 466
QY      224 AlaAlaValGlnGlnTrpLeuGlnLysGlyThrProAlaSerLysLeuIleLeuGlyMet 243
Db      ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
467 TATGCTGTGGGTACTGGATACATAAAGGAATGCCATCAGAGAAGGTGTCATGGGCATC 526
QY      244 ProThrTyrGlyArgSerPheThrLeuAlaSerSerSerAspThrArgValGlyAlaPro 263
Db      ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
527 CCCACATATGGGCACCTCTTCACTATGGCC---TCTGCAGAAACACCACCGTGGGGGCCCT 583
QY      264 AlaThrGlySerGlyThrProGlyProPheThrLysGluGlyMetLeuAlaTrrpTyr 283
Db      ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
584 GCCTCTGGGCCCTGGAGCTCTGACCCCATCACAGAGCTTTCAGGCTTCTGGCCTATTAT 643
QY      284 GluValCysSerTrp---LysGlyAlaThrLysGlnArgIleGlnAspGlnLysValPro 302
Db      ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
644 GAGATCTGCCAGTTCTGAAAGAGAGCAAGATCACGCGGCTCCAGGATCAGCAGGTTCOC 703
QY      303 TyrIlePheArgAspAsnGlnTrpValGlyPheAspAspValGluSerPheLysThrLys 322
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
704 TACGCAGTCAAGGGGAACCAAGTGGGTGGCTATGATGTGAAGAGTATGGAGAACAAAG 763
QY      323 ValSerTyrLeuLysGlnLysGlyLeuGlyGlyValaMetValTrrpAlaLeuAspLeuAsp 342
Db      ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
764 GTTCAGTTCTTAAAGAT-TTAAACCTGGGAGGAGCCATGATCTGGTCTATTGACATGGAT 822
QY      343 AspPheAlaGlyPheSerCysAsnGlnGlyArgTyrProLeuIleGlnThrLeuArgGln 362
Db      ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
823 GACTTCACCTGGCAAAATCCTGCAACAGGGGCCCTTACCTCTTTGTTCAGCAGTCAAGAAAG 882
QY      363 GluLeuSerLeuProTyrLeuPro 370
Db      ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
883 -----CCTTGGTCCCT 894
```

Search completed: June 30, 2003, 01:07:48
Job time : 1478.66 secs

THIS PAGE BLANK (USPTO)

Db 499 FTCPPGTLFDPALHCNW 516

RESULT 10
A33918
chitinase (EC 3.2.1.14) precursor - braconid wasp (*Chelonus* sp.)
C:Species: *Chelonus* sp.
C>Date: 28-Jul-1995 #sequence_revision 28-Jul-1995 #text_change 21-Jul-2000
R:Krishnan, A.; Nair, P.N.; Jones, D.
J. Biol. Chem. 269, 20971-20976, 1994
A:Title: Isolation, cloning, and characterization of new chitinase stored in active form
A:Reference number: A53918; MUID:94342256; PMID:8063715
A:Accession: A53918
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-483 <KRA>
A:Cross-references: GB:U0422; NID:g533504; PIDN:AAA61639.1; PID:g533505
C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 36.1%; Score 864.5; DB 2; Length 483;
Best Local Similarity 37.6%; Pred. No. 1.9e-57;
Matches 178; Conservative 79; Mismatches 171; Indels 45; Gaps 13;

QY 2 KLVCYFTNWAQYRGAEARFLPKDLPSCLTLLIYAFAGMTNHLSTTE-WND-----ETLY 56
DB 23 KVCYFGAWSVYRQNGKFDINGIDPTLCTLLIYSFVGNGDKVKVLDPSDLPGNLDGF 82
QY 57 QEFNGLKKNPKLTKLLAIGGNWFGTKFTDMVATANNRQTFVNSAIRFLRYKYSFDGLDL 116
DB 83 GKFTSLRKNPSVKIMAVGWNAGSVFQSOMASDQATREAFQAVVVKPQQYQDFGDI 142
QY 117 DWEYPGSQSPAVDKERFTTLVQDLANAFQOEAQTSKGERLLLSAAVPAGQTVVDAGYEV 176
DB 143 DWEYPAQRGSGPADVKNWVKLCKALKAFVQ-----HDYLSRAVAAPETASAKSDI 195
QY 177 DKIAQNLDVFNLMAYDFHGSWEKVTGHSNPLYKROESGAAASLVNDAVQOVLQKGTTPA 236
DB 196 AEMSVYLDPLNLTWDFHGFWDGHTGMHAPSSASHSDSGNELKLVNKAAYKWLQNGVPK 255
QY 237 SKLLIGMPTYGRSFTLASSSDTRVGAPATGSGTPGPFYKGGMLAYYEVCSW-KGATQOR 295
DB 256 EKLVGVPAYGKSFSLNSPNSKGLGAPVSGAGTAGPTGNGLLGYNETCEMKGAGDMEV 315
QY 296 IQDOKV-PYIFRDNQWVGDFDVEFKTKVSLKQKGLGAMVWALDLDLDFAGFSCNOGR 353
DB 316 VDQNEKGVPIYAVGQWVSFDDLAIAKAKAQAFIKQEGLGAMVWVETDFFKGL-CGE-K 373
QY 354 YPLIQTLLRQEL-----SLP-----YLPSTGTP-----LEVPKQGPSEPHGSPGQD 396
DB 374 YPVLKALNSVLGRGSSSPAETKRKNVDPDQAPPRSPFAEDSAPEAPVEPVSSESGE- 432
QY 397 TFCQKADGLYNPRRERSFYSC-----AAGRLFOQSCPTGLVFNNSCKCTW 444
DB 433 --CSSVGQPLV---GQNCGLYVCDGDMGGRPKIPQVCPQGLCFNPANNYCDW 480

RESULT 11
A38221
chitinase (EC 3.2.1.14) Mf1 - nematode (*Brugia malayi*)
C:Species: *Brugia malayi*
C>Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jul-2000
R:Fuhrman, J.A.; Lane, W.S.; Smith, R.F.; Plessens, W.F.; Perler, F.B.
Proc. Natl. Acad. Sci. U.S.A. 89, 1548-1552, 1992
A:Title: Transmission-blocking antibodies recognize microfilarial chitinase in brugian
A:Reference number: A38221; MUID:92179220; PMID:1542646
A:Accession: A38221
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid; protein
A:Residues: 1-504 <RUH>
A:Cross-references: GB:M73689; NID:g156063; PIDN:AAA27854.1; PID:g156064
A:Note: sequence extracted from NCBI backbone (NCBIP:85345)

C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 35.9%; Score 860.5; DB 2; Length 504;
Best Local Similarity 37.0%; Pred. No. 4e-57;
Matches 177; Conservative 80; Mismatches 170; Indels 51; Gaps 9;

QY 5 CYFTNWAQYRGAEARFLPKDLPSCLTLLIYAFAGMTNHLSTT---EWNDET-----L 55
DB 27 CYFTNWAQYRGAEARFLPKDLPSCLTLLIYAFAGMTNHLSTT---EWNDET-----L 55
QY 56 YQEFNGLKKNPKLTKLLAIGGNWFGTKFTDMVATANNRQTFVNSAIRFLRYKYSFDGLD 115
DB 85 YSAVTKLRETNPGLKVLVLSYGYGSAIFGIAKSAQKTERFTKSAIAFLRKNFDFGD 144
QY 116 LDWEYPGSQSPAVDKERFTTLVQDLANAFQOEAQTSKGERLLLSAAVPAGQTVVDAGYEV 175
DB 145 LDWEYFVGVCA-----EHAHKLVEAKTAFVEAKTSKQRLLLTAAVSAGKGTIDGSYN 198
QY 176 VDKIAQNLDVFNLMAYDFHGSWEKVTGHSNPLYKROESGAAASLVNDAVQOVLQKGTTP 235
DB 199 VESLGKPNFDLLFLMSYDLHGSWEKNVDLHGKLPHTKGEVSGIGIFNTEFAADYKASKMP 258
QY 236 ASKLLIGMPTYGRSFTLASSSDTRVGAPATGSGTPGPFYKGGMLAYYEVCSW-KGATQ 294
DB 259 KEKIIIGIPMYAQQWTLDPNSETAIGAAASRPSSASKTNPAGGTASYWEICKYLKEGCKE 318
QY 295 RIQDQKV-PYIFRDNQWVGDFDVEFKTKVSLKQKGLGAMVWALDLDLDFAGFSCNOGR 353
DB 319 TVHGEVGAIVKGDQWGYDNEETIRIKMKLKEKYGGAFFIWDLDFDFTGKSCGKGP 378
QY 354 YPLIQTLLRQELSLPYLPSTGTPPELVKPGF-----OPSE 386
DB 379 YPLLNALISELE---GESENFEITTEBSITETAYETDETESETSEAYDTDETEETSE 435
QY 387 PEHGPSQDFTQO---GKADGLYNPRRERSFYSCAAGRLFOQSCPTGLVFNNSCKCC 442
DB 436 TEATYTDTEGEGECPERDGLPHPTDCHLFIQCANNIAYVMQCPATTFPNDIAIKVC 493

RESULT 12
A56596
chitinase (EC 3.2.1.14) - tobacco hornworm
C:Species: *Manduca sexta* (tobacco hornworm)
C>Date: 11-Aug-1995 #sequence_revision 11-Aug-1995 #text_change 29-Jan-1999
R:Kramer, K.J.; Corpuz, L.; Choi, H.K.; Muthukrishnan, S.
Insect Biochem. Mol. Biol. 23, 691-701, 1993
A:Title: Sequence of a cDNA and expression of the gene encoding epidermal and gut chitinase
A:Reference number: A56596; MUID:93357793; PMID:8353525
A:Accession: A56596
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-554 <KRA>
A:Cross-references: GB:U02270; GB:S64757; NID:g406048; PID:g406049
A:Experimental source: larvae
A:Note: sequence extracted from NCBI backbone (NCBIN:136417, NCBIP:136418)
C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 34.4%; Score 824; DB 2; Length 554;
Best Local Similarity 33.7%; Pred. No. 2.6e-54;
Matches 181; Conservative 87; Mismatches 159; Indels 110; Gaps 15;

QY 1 AKLVCFYTNWAQYRGAEARFLPKDLPSCLTLLIYAFAGMTNHLSTTEWMDTEL----- 55
DB 23 ARIVCFYNWAVYRQVRYGIEDIPVEKCTHIIYSFIGV-----TEGSEVLIIDPE 75
QY 56 -----YQEFNGLKKNPKLTKLLAIGGNWFGTKFTDMVATANNRQTFVNSAIRFLRK 108
DB 76 LDVDKNGFRNFTSLRSSHPVSKFWAVGVGWAEGSSKYSHMVAQKSTRMSFIRSVSFLKK 135
QY 109 YSFDGLDLDEYVPS--OGSPAVDKERFTTLVQDLANAFQOEAQTSKGERLLLSAAVPAG 166
DB 136 YDFDGLDLDEYVPGAAADRGSGFSKDKFLYLVLQELRRAPFIRVGK--GWE---LTAAPLA 190

```
QY 167 QTVYDAGYEVDKIAQNLDFVNLMAVDYFHGSWEKVTGHNPSLYKROBESGAASLNVDAAV 226
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 191 NFRLEMEGYHVPCLCOELDAIHVMSYDLRGWAGFADVHSPLYKRPDQWAEKLNVDGL 250

QY 227 QOWLQGTSPASKLILGMPYGRSFTLASSD-----TRVGAPATGSGTGGPFTKEGMLA 281
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 251 HLWEEKGCPNKLWVGIPFYGRSFTLSAGNNYVGLGTFINKEA-GGGDPAPYTNATGFWA 309

QY 282 YVEVCSW-----KGATKQRIQODQKVPYIFRDQWVGFDDVESFKTVSVLKOKGLGGAMV 336
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 310 YVEICTEVKDQSGWTKWDEQCKPYAYKGTQWVGEDQPSVEIKWIKOKGYLGAMT 369

QY 337 WALDLDDFAGFSCNCGRYPLIOTLRQELSLYLP-----SGTPELEVPKQCPSEPSH 389
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 370 WALDDDDFOGL-CGE-KNPLIKILHKHMSYVPPPHTEHTPTPEWARP-FSTPSDPSE 426

QY 390 G-----PSPGQDT--- 397
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 427 GDIPIPTTTAKPASTTKTTVKTITTTTAKPQSVDEENDINVRPEKPEPEPEVEVP 486

QY 398 -----FCQKADGLYPNPRSSYFSCAAGRLFOQSCPTGLVFSNCKCTM 444
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 487 PTNEVDGSEICNSDQDYI-PDKKHCDKYRWCNGEAMQFSCQHGTVFVNLVCDW 542

RESULT 13
T14075
Chitinase (EC 3.2.1.14) - yellow fever mosquito
C:Species: Aedes aegypti (yellow fever mosquito)
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T14075
R:de la Vega, H.; Specht, C.A.; Liu, Y.; Robbins, P.W.
Insect Mol. Biol. 7, 233-239, 1997
A:Title: Chitinases are a multi-gene family in Aedes, Anopheles, and Drosophila.
A:Reference number: Z17872
A:Accession: T14075
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1635 <DEL>
A:Cross-references: EMBL:AF026492; NID:g2564720; PID:g2564721; PIDN:AA81850.1
C:Genetics:
A:Gene: CHT2
A:Introns: 462/3; 524/3; 618/1; 951/3; 1151/2
C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 32.8%; Score 787.5; DB 2; Length 1635;
Best Local Similarity 41.4%; Pred. NO. 7e-51;
Matches 157; Conservative 72; Mismatches 121; Indels 29; Gaps 11;

QY 2 KLVCYFTNWAQYRQGEARFLPKDLPSLCTHLYAFAGTNNHQLST---TEWN--DETLY 56
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1258 KVCYFTNWAHVRQGGKVTPTDDIDSNLCTHIVGFAVLDRSLTIKTHDSWADINQFY 1317

QY 57 QFNGLKQKNPKLKTLLAIGGNWFGT-QKFTDMVATANNRQTFVNSAIRFLRYKISYDGLD 115
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1318 ERVVEHKRGTKV--TLALGGWNSLGDYKSVKLVRSSARRAFVXIAVEFIEKYNFDGLD 1375

QY 116 LDWEYP-----GSQSPAVDKERFTTHVQDLANARQGEAQTSGKERLLISAAVPAGTY 169
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1376 LDWEYFVQVQVDCCKGFPD-EKEGFAELVKELAEFR-----PRKLLLSAAVSP$KXV 1427

QY 170 VDAGYEVDKIAQNLDFVNLMAVDYFHGSWEKVTGHNPSLYKROBESGAASLNVDAAVQW 229
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1428 IDAGYDVPVLAIEYFDVIAWMTYDFHGHNDKQTHGVAFLYYPGDT--YDYFNANFSINTW 1485

QY 230 LQKGTSPASKLILGMPYGRSFTLASSDTRVGAPATGSGTGGPFTKEGMLAYEVC--- 286
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1486 IEKGAPSRKLVMGMPYLGQSFSLSADSGKNGLNDKSYGPGEGAGQFTRAGGFLAFYEICEKV 1545

QY 287 SWGATKQRIQODQKVPYIFRDQWVGFDDVESFKTVSVLKOKGLGGAMVWALDDDFEA 345
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1546 NQNGWAVRDSGRIGFYALRGQWVSVDVDEIRRSQFVKKNNLGGGMIWALDDDFR 1605
```

```
QY 346 GFSCNCGRYPLIOTLRQEL 364
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1606 G-RCGCGRHPLLKTMNQEL 1623

RESULT 14
A38368
Chitinase (EC 3.2.1.14) precursor - Bacillus circulans
C:Species: Bacillus circulans
C>Date: 28-Jun-1991 #sequence_revision 28-Jun-1991 #text_change 15-Oct-1999
C:Accession: A38368
R:Watanabe, T.; Suzuki, K.; Oyanagi, W.; Ohnishi, K.; Tanaka, H.
J. Biol. Chem. 265, 15659-15665, 1990
A:Title: Gene cloning of chitinase A1 from Bacillus circulans WL-12 revealed its evolutic
A:Reference number: A38368; MUID:90368776; PMID:2203782
A:Accession: A38368
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-699 <WAT>
A:Cross-references: GB:M57601; GB:J05599; NID:g1066341; PIDN:AAA81528.1; PID:g142688
C:Superfamily: fibronectin type III repeat homology
C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 24.4%; Score 584.5; DB 2; Length 699;
Best Local Similarity 33.0%; Pred. NO. 4.5e-36;
Matches 147; Conservative 71; Mismatches 143; Indels 85; Gaps 16;

QY 2 KLVCYFTNWAQYRQGEARFLPKDLPSLCTHLYAFAGM-----ETLYQ-----EFNGLKKNPKLKTLLAI 75
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 45 KLVGYPSMAAYGR---NYNVADIDPTKVTHTINAFADICWNGIHGNPDGPNPVTWTC 101

QY 41 TNHQLSTTE-----WND-----ETLYQ-----EFNGLKKNPKLKTLLAI 75
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 102 QNEKQSTINVPNGTIVLGDPIWIDTGKTFAGTWDQPIAGNINQLNKLQTNENLKTIIISV 161

QY 76 GGNFQGTOKFTDMVATANNRQTFVNSAIRFLRYKISYDGLDWEYPSGQ-----SPAVD 130
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 162 GGTW-SNRFSDVAATAATREVFAANSVDFLRKYNFDFGVLDWEYFVSGGLDGNKRPE 220

QY 131 KERFTTLVQDLANARQGEAQTSGKERLLISAAVPAGTYVDAGYEVDKIAQNLDFVNLMA 190
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 221 KONYTLLSKIREKLDAAAGAVDGK-KYLLTIASGASATYA-ANTELAIAIVDWINMT 278

QY 191 YDFHGSWEKVTGHNPSLYKROBESGA-----AASLNVDAAVQWLOKQTPASKLILGMPY 246
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 279 YDFNGAWQKISAHNAPLNYDPAASAAGVDPDANTFNVAAGAQLDAGVPAKLVGVFPY 338

QY 247 GRSFTLASSDTRVGAPATGSGTGGPFTKEGMLAYEV---CSWKGATKQRIQODQKVP 302
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 339 GRGWDGCAQAGNGQYQTCCTGGSSVG--TWEAGSFDYDLEANYINKNGYTRYWINDTAKVP 396

QY 303 YIFR--DNQWVGFDDVESFKTVSVLKOKGLGGAMVWALDDDFAGFSCNCGRYPLIOTL 360
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 397 YLYNASNKRFTSYDDAESVGYKTAYIKSGLGGAMFWELSGD-----RNKTL 443

QY 361 ROELSLPLPSG--TPELEVPKPGQP 384
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 444 QNKLKAD-LPTGTVPPVDTTAPSV 468

RESULT 15
D83764
Chitinase BH0916 [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: D83764
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, F.; Hiran
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: D83764
A:Status: preliminary
```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2003, 20:51:08 ; Search time 7.89941 Seconds
(without alignments)
2336.500 Million cell updates/sec

Title: US-10-004-219B-10

Perfect score: 2398

Sequence: 1 AKLVCFYFTNWAQYRGAEARF.....QQSCPTGLVFSNCKCKCTWN 445

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1045.5	43.6	390	1 C312 HUMAN	Q15782 homo sapien
2	1041.5	43.4	383	1 C311 HUMAN	P36222 mus sapien
3	984.5	41.1	381	1 C311 MOUSE	Q61362 mus musculus
4	964.5	40.2	537	1 OGP BOVIN	Q28042 bos taurus
5	959.5	40.0	539	1 OGP SHEEP	Q28542 ovies aries
6	956	39.9	617	1 CHIT CAEL	Q11174 caenorhabdi
7	947.5	39.5	527	1 OGP FIG	Q28990 sus scrofa
8	942	39.3	721	1 OGP MOUSE	Q62010 mus musculus
9	927	38.7	671	1 OGP MESAU	Q60557 mesocricetu
10	925.5	38.6	623	1 OGP PAPAN	P36718 papio anubi
11	917.5	38.3	678	1 OGP HUMAN	Q12889 homo sapien
12	860.5	35.9	504	1 CHIT BRUMA	P29030 brugia mala
13	824	34.4	554	1 CHIT MANSE	P36362 manduca sex
14	584.5	24.4	699	1 CHIT BACCI	P20533 bacillus ci
15	493.5	20.6	423	1 CH11 APHAL	P32470 aphanoclad
16	478	19.9	423	1 CH14 TRIHA	P48827 trichoderma
17	425.5	17.7	427	1 CH11 COCIM	P54196 coccidioid
18	422	17.6	499	1 CH1B SERMA	P11797 serratia ma
19	357	14.9	550	1 CHIT NPVOP	P07254 orgyia pseu
20	355.5	14.8	563	1 CH1A SERMA	P36909 streptomyce
21	331.5	13.8	619	1 CHIT STRLI	P12220 streptomyce
22	331.5	13.8	820	1 CH1A ALTOS	P32823 alteromonas
23	329	13.7	551	1 CHIT NPVAC	P41684 autographa
24	328	13.7	1046	1 CH1D VIBFU	P96156 vibrio furn
25	305	12.7	610	1 CHIT STRPL	P09805 kluyveromyc
26	270.5	11.3	1146	1 KTXA KUJLA	P01460 rattus norv
27	161	6.7	367	1 DIAC RAT	P37531 bacillus su
28	155	6.5	427	1 YAAH BACSU	Q01459 homo sapien
29	147.5	6.2	385	1 DIAC HUMAN	Q14599 caenorhabdi
30	127.5	5.3	491	1 YKCS CAEL	Q07833 bacillus su
31	122	5.1	2334	1 WAPA BACSU	P27050 bacillus ci
32	120.5	5.0	524	1 CH1D BACCI	P23989 streptococ
33	113	4.7	1026	1 BGAL STRTR	

RESULT 1
C312 HUMAN
ID C312 HUMAN STANDARD; PRT; 390 AA.
AC Q15782; Q15783; Q15749;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Chitinase 3-like protein 2 precursor (YKL-39) (Chondrocyte protein 39).
DE CH13L2.
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (SHORT AND LONG ISOFORMS).
RA Grossman A., Matsuyama T., Baker E., Waterhouse P., Sutherland G.R., Mak.F.W.;
RT "Cloning of a novel lymphoid restricted human chitinase and localization to lp13.3".
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.

SEQUENCE FROM N.A. (SHORT ISOFORM).
TISSUE=Articular cartilage;
MEDLINE=96325055; PubMed=8702629;
Hu B., Trinh K., Figueira W.F., Price P.A.;
"Isolation and sequence of a novel human chondrocyte protein related to mammalian members of the chitinase protein family.";
J. Biol. Chem. 271:19415-19420(1996).
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A SHORT FORM (SHOWN HERE) AND A LONG FORM; MAY BE PRODUCED BY ALTERNATIVE SPLICING.
CC -!- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN CHONDROCYTES, FOLLOWED BY SYNOVIOCYTES, LUNG AND HEART. NOT DETECTED IN BRAIN, SPLEEN, PANCREAS, AND LIVER.
CC -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announcement/ or send an email to license@isb-sib.ch).
EMBL; U58515; AAB04534.1; -;
EMBL; U58514; AAB04533.1; -;
EMBL; U49835; AAC50597.1; ALT_INIT.
MIM; 601526; -;
InterPro; IPR001579; Chitinase 18/2.
InterPro; IPR001223; Glyco_hydro_18.
Pfam; PF00704; Glyco_hydro_18; 1.
ProDom; PD000471; Glyco_hydro_18; 1.
PROSITE; PS01095; CHITINASE_18; FALSE NEG.
KW Hydrolase; Signal; Alternative splicing.
FT SIGNAL 1 26
POTENTIAL.

ALIGNMENTS


```

Db      376 IKDALA 381

RESULT 3
C3L1 MOUSE
ID C3L1 MOUSE STANDARD; PRT; 381 AA.
AC Q61362;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Chitinase-3 like protein 1 precursor (Cartilage glycoprotein-39)
DE (GP-39) (BRP39 protein).
GN CH13L1 OR BRP39.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Breast;
RX MEDLINE=95060797; PubMed=7970700;
RA Morrison B.W., Leder P.;
RT "neu and ras initiate murine mammary tumors that share genetic markers
RT generally absent in c-myc and int-2-initiated tumors.";
RL Oncogene 9:3417-3426(1994).
CC -!- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN THE CAPACITY OF CELLS TO
CC RESPOND TO AND COPE WITH CHANGES IN THEIR ENVIRONMENT (BY
CC SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Extracellular (By similarity).
CC -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X93035; CAA63603.1; -
DR MGD; MGI:1340899; Ch13l1.
DR InterPro; IPR001579; Chitinase_18/2.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
KW Glycoprotein; Signal.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 381
FT CARBOHYD 60 630 N-LINKED (GLCNAC... ) (POTENTIAL).
SQ SEQUENCE 381 AA; 43001 MW; EF6581E8184F0450 CRC64;

Query Match 41.1%; Score 984.5; DB 1; Length 381;
Best Local Similarity 50.5%; Pred. No. 1.6e-67;
Matches 185; Conservative 73; Mismatches 99; Indels 9; Gaps 5;

QY 2 KLVCFYFTNAQVROGEARFLPKDLPDLCTHLIYAFAGM-TNHQLSTTWENDE-TLYQEFN 60
Db 23 KLVCFYFTNAQVROGEARFLPKDLPDLCTHLIYAFAGM-TNHQLSTTWENDE-TLYQEFN 82
QY 61 GLKXNPKLTLAIGGWNFGTKFTDMVATANNROTFFVNSAIRFLRYKYSFGDGLDMEY 120
Db 83 KLKTRNTNLKLLSVGGWKEGKEKFEISANTERRTAFVRSVAPFLRSYGFGLDLAWLY 142
QY 121 PQSQSPAVDKERFTTLVDLANAFOEAQATGKERLLLSAAVAPGQTVYDAGYEVDKTA 180
Db 143 PLRL-----DKQYFSTFLIKELNAEFTKEVQ-PGREKLLLSAALSAGKVAIDTGYDIAQIA 196
QY 181 ONLDFVNLMAYDFHGSWEKVTGHNSPLYKRESGAASLNDVAQVQLQGTGTPASKLI 240
Db 197 QHLDFTNLMTYDFHGWROITGHHSPLFOGQKDFRDYSNVNYAVQYIMRILGAQASKLL 256
QY 241 LGMPTYGRSFTLASSSDTRVGAPATGSGTFFGFTKEGGMLAYVEVCSW-KGATKQRIQDQ 299

Db      257 MGIPTFGKSFTLA-SSENQLGAPISGEGLPGRFTKEAGTFLAYEICDFELKGAEVHRLSNE 315
QY 300 KVPYIFRDNQWVGPDVSEFKTKVSLYKQKGGWVWALDLDLDFAGSCNQGRYPLIQT 359
Db 316 KVPFATKGNQWVGVEHKESVKNKVGFLKEKLAGANWVWALDLDLDFQGTGTCQPKRFPPLTNA 375
QY 360 LRQELS 365
Db 376 IKDALA 381

RESULT 4
OGP BOVIN
ID OGP BOVIN STANDARD; PRT; 537 AA.
AC Q28042;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Oviduct-specific glycoprotein precursor (Oviductal glycoprotein)
DE (Oviductin) (Estrogen-dependent oviduct protein) (Fragment).
GN OVGp1 OR OGP.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
SEQUENCE FROM N.A., AND SEQUENCE OF 19-47.
RC TISSUE=Oviduct;
RX MEDLINE=942527768; PubMed=8199272;
RA Sendai Y., Abe H., Kikuchi M., Satoh T., Hoshi H.;
RT "Purification and molecular cloning of bovine oviduct-specific
RT glycoprotein.";
RL Biol. Reprod. 50:927-934(1994).
CC -!- FUNCTION: BINDS TO OOCYTE ZONA PELLUCIDA IN VIVO. MAY PLAY A ROLE
CC IN THE FERTILIZATION PROCESS AND/OR EARLY EMBRYONIC DEVELOPMENT.
CC -!- SUBCELLULAR LOCATION: Secretory granules.
CC -!- TISSUE SPECIFICITY: OVIDUCT.
CC -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D16639; BAA04065.1; -
DR InterPro; IPR001579; Chitinase_18/2.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
KW Glycoprotein; Fertilization; Signal.
FT NON_TER 1 18
FT SIGNAL <1 18
FT CHAIN 19 537
FT CARBOHYD 399 399 N-LINKED (GLCNAC... ) (POTENTIAL).
SQ SEQUENCE 537 AA; 59617 MW; CFCBEB6F021D791 CRC64;

Query Match 40.2%; Score 964.5; DB 1; Length 537;
Best Local Similarity 51.0%; Pred. No. 8.5e-66;
Matches 186; Conservative 60; Mismatches 112; Indels 7; Gaps 3;

QY 2 KLVCFYFTNAQVROGEARFLPKDLPDLCTHLIYAFAGM-TNHQLSTTWENDE-TLYQEFN 60
Db 20 KLVCFYFTNAQVROGEARFLPKDLPDLCTHLIYAFAGM-TNHQLSTTWENDE-TLYQEFN 79
QY 61 GLKXNPKLTLAIGGWNFGTKFTDMVATANNROTFFVNSAIRFLRYKYSFGDGLDMEY 120
Db 80 KLKERNRGLKLLSIGGWNFGTVTRFTMLSTFNSRERFVSSVIALLRTHGFGDGLDFFLY 139

```



```
QY 121 PGSGSPAVDKERFTTLVODLANAFOEAOCTSGKERLLLSAAVPAQTYVDAGYEVDKIA 180
DB 140 PGLRSGSPARDRTWTFVLELLQAFKNEAQLTWPRLLLSAAVSGDPHVVKAYEARLLG 199
QY 181 QNLDFVNLMAIDFHGSEKVTGHNPSLYKROBESGAASLNDAVQOVLQKGTTPASKLI 240
DB 200 RLDDFISVLSYDLHGSWEKVTGHNPSLPLGPKSSA-----YAMNYWRLQGVPEKLL 254
QY 241 LGNPTVGRSFTLASSSDTRVGAATSGTPGPTKEGGMLAYVEVCSW-KGATKQRIQOQ 299
DB 255 MGLPTVGRTPHLLKASQNEURAQVGPASPGKTYKQAGFLAYYEICCFVRRAKKRINDQ 314
QY 300 KVPYIFRDNQWGFDDVSEFPTKVSYLKOKGLGAMVWALDDDFAGFCSCNQGRIPLIOT 359
DB 315 YVPYAFKGEWGYDDAISFGYKAFIKREHFGAMVWTLDDDFRGYFCGTGPFPLVHT 374
QY 360 LRQEL 364
DB 375 LNNLL 379

RESULT 5
OGP SHEEP
ID OGP SHEEP STANDARD; PRT; 539 AA.
AC Q28542; Q28543;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE Oviduct-specific glycoprotein precursor (Oviductal glycoprotein)
DE (Oviductin) (Estrogen-dependent oviduct protein) (Estrus-associated
DE oviductal glycoprotein) (OEGP).
GN OVGPI OR OGP.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 22-39.
RC TISSUE=Oviduct;
RX MEDLINE=95269691; PubMed=7750470;
RA Desouza M.M., Murray M.K.;
RT "An estrogen-dependent secretory protein, which shares identity with
RT chitinase, is expressed in a temporally and regionally specific
RT manner in the sheep oviduct at the time of fertilization and embryo
RT development."
RL Endocrinology 136:2485-2496(1995).
RN [2]
RP SEQUENCE OF 10-539 FROM N.A.
RC STRAIN=Merino; TISSUE=Oviduct;
RX MEDLINE=96329120; PubMed=8726871;
RA Marshall J.T.A., Nancarrow C.D., Brownlee A.G.;
RT "Cloning and sequencing of a cDNA encoding an ovine
RT oestrus-associated oviductal protein."
RL Reprod. Fert. Dev. 8:305-310(1996).
CC -1- FUNCTION: BINDS TO OOCYTE ZONA PELLUCIDA IN VIVO. MAY PLAY A ROLE
CC IN THE FERTILIZATION PROCESS AND/OR EARLY EMBRYONIC DEVELOPMENT.
CC -1- SUBCELLULAR LOCATION: Secretory granules.
CC -1- TISSUE SPECIFICITY: OVIDUCT.
CC -1- DEVELOPMENTAL STAGE: LEVELS ARE HIGHEST IN THE FIMBRIA AND AMPULLA
CC AT ESTRUS AND ON DAY 1 OF PREGNANCY, WHEN GAMETE TRANSPORT AND
CC FERTILIZATION OCCURS IN THE E2-DOMINATED FALLOPIAN TUBE. LEVELS
CC DECLINE SIGNIFICANTLY ON DAY 2 AND UNDERGO A FURTHER SIGNIFICANT
CC REDUCTION ON DAY 3 OF PREGNANCY COINCIDENT WITH TRANSPORT OF THE
CC EMBRYO FROM THE OVIDUCT TO THE UTERUS, A REPRODUCTIVE STAGE
CC ASSOCIATED WITH RISING PROGESTERONE LEVELS.
CC -1- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
```

```
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; U16719; AAC48471.1; -
DR EMBL; U17988; AAB01052.1; -
DR InterPro; IPR001579; Chitinase_18/2.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF00704; Glyco_hydro_18; 1
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
KW Glycoprotein; Fertilization; Signal.
FT SIGNAL 1 21 BY SIMILARITY.
FT CHAIN 22 539 OVIDUCT-SPECIFIC GLYCOPROTEIN.
FT CARBOHYD 402 402 N-LINKED (GLCNAC... ) (POTENTIAL).
FT CONFLICT 13 13 M -> V (IN REF. 2).
FT CONFLICT 122 122 K -> N (IN REF. 2).
FT CONFLICT 282 282 A -> V (IN REF. 2).
FT CONFLICT 375 375 R -> V (IN REF. 2).
FT CONFLICT 484 484 A -> H (IN REF. 2).
FT CONFLICT 520 520 I -> T (IN REF. 2).
SQ SEQUENCE 539 AA; 59535 MW; F35000269987C193 CRC64;

Query Match 40.0%; Score 959.5; DB 1; Length 539;
Best Local Similarity 51.0%; Pred. No. 2e-65;
Matches 186; Conservative 57; Mismatches 115; Indels 7; Gaps 3;

QY 2 KLVCVFTWVAQVROGEARFLPKDLPDLCTHLIYAPAGTWNHOLSTTEWDE-TLYQREN 60
DB 23 KLVCVFTWVAQVROGEARFLPKDLPDLCTHLIYAPAGTWNHOLSTTEWDE-TLYQREN 82
QY 61 GLKKNPKLTLAIGGNWFGTKETDMVATANNRQTFVNSAIRFLRYSPDGLDLDMWEY 120
DB 83 KLVKRNGLKTLVSVGGWNGFSTRFKMLSTFSNERFVKSVIALLRTHGFGDLDFLY 142
QY 121 PGSGSPAVDKERFTTLVODLANAFOEAOCTSGKERLLLSAAVPAQTYVDAGYEVDKIA 180
DB 143 PGLRSGSPARDRTWTFVLELLQAFKNEAQLTWPRLLLSAAVSGDPHVVKAYEARLLG 202
QY 181 QNLDFVNLMAIDFHGSEKVTGHNPSLYKROBESGAASLNDAVQOVLQKGTTPASKLI 240
DB 203 RLDDFISVLSYDLHGSWEKVTGHNPSLPLGPKSSA-----YAMNYWRLQGVPEKLL 257
QY 241 LGNPTVGRSFTLASSSDTRVGAATSGTPGPTKEGGMLAYVEVCSWKATQKQ-IOQ 299
DB 258 MGLPTVGRTPHLLKASQNEURAQVGPASPGKTYKQAGFLAYYEVCVFQRAKRWINDQ 317
QY 300 KVPYIFRDNQWGFDDVSEFPTKVSYLKOKGLGAMVWALDDDFAGFCSCNQGRIPLIOT 359
DB 318 YVPYAFKGEWGYDDAISFGYKAFIKREHFGAMVWTLDDDFRGYFCGTGPFPLAHT 377
QY 360 LRQEL 364
DB 378 LNNLL 382

RESULT 6
CHIT_CABEL STANDARD; PRT; 617 AA.
ID CHIT_CABEL
AC Q11174;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Putative endochitinase (EC 3.2.1.14).
GN C04F6.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Nhan M.;
```

Submitted (DEC-1995) to the ENBL/GenBank/DBJ databases.
 -!- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages of N-acetyl-D-glucosamine polymers of chitin.
 -!- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL HYDROLASES).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

DR	ENBL; U42835; AAA83586.1; -.
DR	WormPep; C04F6.3; CE03923.
DR	InterPro; IPR002557; Chitin_bind_PeraA.
DR	InterPro; IPR001579; Chitinase_18/2.
DR	InterPro; IPR001223; Glyco_hydro_18.
DR	Pfam; PF00704; Glyco_hydro_18; 1.
DR	Pfam; PF01607; CBM_14; 2.
DR	ProDom; PD000471; Glyco_hydro_18; 1.
DR	SMART; SM00494; CheBD2; 2.
DR	PROSITE; PS01095; CHITINASE_18; 1.
KW	Hypothetical protein; Hydrolase; Glycosidase; Chitin degradation.
FT	ACT_SITE 179 / 179 PROTON DONOR (BY SIMILARITY).
FT	DOMAIN 435 / 478 THR-RICH.
SQ	SEQUENCE 617 AA; 66857 MW; DDALD2AAAC0E54DA CRC64;

RESULT 7

OGP_PIG	OGP_PIG	STANDARD;	PRT;
ID	Q28590;		527 AA.
AC			
DT	01-NOV-1997	(Rel. 35, Created)	
DT	01-NOV-1997	(Rel. 35, Last sequence update)	
DT	15-JUN-2002	(Rel. 41, Last annotation update)	

```
Db 378 LNDLL 382
RESULT 8
ID OGP MOUSE STANDARD; PRT; 721 AA.
AC Q62010;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Oviduct-specific glycoprotein precursor (Oviductal glycoprotein)
DE (Oviductin) (Estragen-dependent oviduct protein).
GN OVGPI OR OGP OR CHITS.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=ICR; TISSUE=Oviduct;
RX MEDLINE=96115001; PubMed=7492680;
RA Sendai Y., Komiya H., Suzuki K., Onuma T., Kikuchi M., Hoshi H.,
Araki Y.;
RT "Molecular cloning and characterization of a mouse oviduct-specific
glycoprotein.";
RL Biol. Reprod. 53:285-294(1995).
CC -!- FUNCTION: BINDS TO OOCYTE ZONA PELLUCIDA IN VIVO. MAY PLAY A ROLE
IN THE FERTILIZATION PROCESS AND/OR EARLY EMBRYONIC DEVELOPMENT.
CC -!- SUBCELLULAR LOCATION: Secretory granules.
CC -!- TISSUE SPECIFICITY: EPITHELIAL CELLS OF THE OVIDUCT.
CC -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC EMBL; D32137; BAA06863.1; -.
DR MGD; MGI:106661; Ovgpl.
DR InterPro; IPR001579; Chitinase_18/2.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR PROSITE; PS01095; CHITINASE_18; FALSE NEG.
KW Glycoprotein; Fertilization; Repeat; Signal.
FT SIGNAL 1 21
FT CHAIN 22 721
FT DOMAIN 486 632
FT [TAP]-G-[IV]
FT CARBOHYD 402 402 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 442 442 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 469 469 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 721 AA; 78807 MW; 37246C8F01665652 CRC64;
Query Match 39.3%; Score 942; DB 1; Length 721;
Best Local Similarity 49.3%; Pred. No. 6.5e-64;
Matches 185; Conservative 60; Mismatches 122; Indels 8; Gaps 4;
QY 2 KLVCFYTNWAQYRQGEARFLPDKDPSLCTHLTYAFAGMTHQSLSTEWNDE-TLYQEEN 60
Db 23 KLVCFYTNWAHSRPGASIMPHDDLPCLTHLIFAFASHNNQIVAKNQDENLVLYPEEN 82
QY 61 GLKMMPKLTLIAIGGNWFGTKQFTDMVATANNROTFFNSAIRFKYKVSFGDLIDWEY 120
Db 83 KLKERNRELKTLISIGGNWFGTSRFTAMLSLTANREKFDVSVISFLRFGFDGLDFFLY 142
QY 121 PGSGSPAVDKERFTTLVQDLANAFQEAQTSCKERLLLSAAVPAGQTVYDAGEYDKIA 180
Db 143 PLGRGSPPHDWNFLFIEBELQAFREALLTOHPRLLSAAVSGIPSIHTSYDALLIG 202
```

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; D32218; BAA06977.1; -
DR EMBL; U15048; AAC53584.1; -
DR InterPro; IPR001579; Chitinase 18/2.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR PROSITE; PS01095; CHITINASE 18; FALSE NEG.
KW Glycoprotein; Fertilization; Repeat; Signal.
FT SIGNAL 1 21 OVIDUCT-SPECIFIC GLYCOPROTEIN.
FT CHAIN 22 671 8 X 15 AA TANDEM REPEATS.
FT DOMAIN 490 609 1.
FT REPEAT 490 504 2.
FT REPEAT 505 519 3.
FT REPEAT 520 534 4.
FT REPEAT 535 549 5.
FT REPEAT 550 564 6.
FT REPEAT 565 579 7.
FT REPEAT 580 594 8.
FT REPEAT 595 609 9.
FT CARBOHYD 402 402 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 511 511 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 526 526 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 541 541 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 556 556 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 571 571 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 586 586 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 26 26 H -> A (IN REF. 4).
FT CONFLICT 33 33 H -> I (IN REF. 4).
FT CONFLICT 137 137 R -> G (IN REF. 2 AND 3).
FT CONFLICT 153 153 R -> Q (IN REF. 2 AND 3).
FT CONFLICT 165 165 F -> Y (IN REF. 2 AND 3).
FT CONFLICT 193 193 Q -> L (IN REF. 2 AND 3).
FT CONFLICT 531 545 MISSING (IN REF. 2 AND 3).
FT CONFLICT 595 595 T -> I (IN REF. 2 AND 3).
SQ SEQUENCE 671 AA; 73250 MW; BB57E0E514EC1972 CRC64;

Query Match 38.7%; Score 927; DB 1; Length 671;
Best Local Similarity 46.9%; Pred. No. 8.1e-63;
Matches 184; Conservative 61; Mismatches 135; Indels 12; Gaps 4;

QY 2 KLVCFYFTWNAOYRQGEARFLPKDLPDLPSLCTHLIYAFAGTWNHOLSTTEWDE-TLYQSFN 60
DB 23 KLVCFYFTWNAHSRPGPASILPRDLDPFLCTHLIFAFASNNQIVANNLQDEKILYPFN 82
QY 61 GLKKNPKLKTLLAIGGWNFGTKFTDMVATANNROTFTVNSAIRFLRKYSPDGLDWEY 120
DB 83 KLKERNRALKTLISVGGWNGFTSRFTTLMSTLASREKFGSVSLRTHGFDGLDFLY 142
QY 121 PEGSQSPAVDKERFTTLVODLANAQOBAQTSKGKRLLSAAVPAGQTVYDAGYEVKIA 180
DB 143 PGLRGSPINDRNFLFLBELQFAFEKALLQRPRLLSAAVSGIPYIQTSDVHLIG 202
QY 181 QNLDFVNLWAYDFHGSWEKVTGHNPSFLYKQESGAASLNVDAVQWLQGTASKLI 240
DB 203 RRLDPIVNLVSYDLHGSWEKSTGHNPSFLPDPKSSA-----FAMNYRNILGAPADKIL 257
QY 241 LGMPYGRSFTLASSSDTRVGPATGSGTPGFTKEGGMLAYEYVCSW-KGATKQRIQDQ 299
DB 258 MGFPAVGRFTLLRRESKGLQAASGWPASPGYTKQAGFLAYEYVCSFIQRAEKWIHQ 317
QY 300 KVPYIFRDNQWGFDDVFSFKTKVSYLKQKGLGGAMWALDLDPPAGFSCNGRYPLIQT 359
DB 318 YVPYAYKGEWGYDDAVSFSYKAMFVKEHFGGAMVTLDMDDVRGTCGNGPFPPLVHI 377
QY 360 L-----RQELSLPLPSGTPELEVKPGQSP 386
DB 378 LNELLVRAEFNSTPLPQFQWFTLLPVNSSGPGSE 409

RESULT 10
OGP PAPAN
ID- OGP PAPAN STANDARD; PRT; 623 AA.
AC P36718;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Oviduct-specific glycoprotein precursor (Oviductal glycoprotein)
DE (Oviductin) (Estrogen-dependent oviduct protein).
GN OVEP1 OR OGP.
OS Papio anubis (Olive baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC Cercopithecinae; Papio.
OX NCBI_taxID=9555;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Oviduct;
RX MEDLINE=98244335; PubMed=9584944;
RA Verhage H.G., Fazleabas A.T., Mavrogianis P.A., O'Day-Bowman M.B.,
RA Donnelly K.M., Arias E.B., Jaffe R.C.;
RT "The baboon oviduct: characteristics of an oestradiol-dependent
RT oviduct-specific glycoprotein.";
RL Hum. Reprod. Update 3:541-552(1997).
RN [2]
RP PRELIMINARY PARTIAL SEQUENCE FROM N.A.
RC TISSUE=Oviduct;
RX MEDLINE=91367180; PubMed=1716345;
RA Donnelly K.M., Fazleabas A.T., Verhage H.G., Mavrogianis P.A.,
RA Jaffe R.C.;
RT "Cloning of a recombinant complementary DNA to a baboon (Papio
RT anubis) estradiol-dependent oviduct-specific glycoprotein.";
RL Mol. Endocrinol. 5:356-364(1991).
CC -1- FUNCTION: BINDS TO OOCYTE ZONA PELLUCIDA IN VIVO. MAY PLAY A ROLE
CC IN THE FERTILIZATION PROCESS AND/OR EARLY EMBRYONIC DEVELOPMENT.
CC -1- SUBCELLULAR LOCATION: Secretory granules.
CC -1- TISSUE SPECIFICITY: OVIDUCT.
CC -1- DEVELOPMENTAL STAGE: AT THE TIME OF OVULATION.
CC -1- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; M59903; AAB39765.1; -
DR PIR; A37954; A37954.
DR InterPro; IPR001579; Chitinase 18/2.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR PROSITE; PS01095; CHITINASE 18; FALSE_NEG.
KW Glycoprotein; Fertilization; Signal.
FT SIGNAL 1 21 BY SIMILARITY.
FT CHAIN 22 623 OVIDUCT-SPECIFIC GLYCOPROTEIN.
FT CARBOHYD 402 402 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 441 441 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 623 AA; 69291 MW; 9B21CB481FF1268 CRC64;

Query Match 38.6%; Score 925.5; DB 1; Length 623;
Best Local Similarity 44.1%; Pred. No. 9.6e-63;
Matches 194; Conservative 66; Mismatches 135; Indels 45; Gaps 8;

QY 2 KLVCFYFTWNAOYRQGEARFLPKDLPDLPSLCTHLIYAFAGTWNHOLSTTEWDE-TLYQSFN 60
DB 23 KLVCFYFTWNAHSRPGPASILPRDLDPFLCTHLIFAFASNNQIVAKLODEKILYPFN 82
QY 61 GLKKNPKLKTLLAIGGWNFGTKFTDMVATANNROTFTVNSAIRFLRKYSPDGLDWEY 120

Db 83 KLKERNRELKTLISIGGNFGTSRFTTLMSTFANREKFIASVLSLRTHDFDGLDLFFLY 142
QY 121 PGSGGSPAVDKERFTTLVODLANAFOEAQTSKGERLLLSAAVPAGQTYVDAGYVDKIA 180
Db 143 PGLRGSPMDHRTWFLFLIEELLFAFRKEALLTWPRLLLSAAVGVPHIVQTSYDVRFLG 202
QY 181 QNLDFVNLMAYPFHGWEKVTGHNSPLYKQESGAAASLVNDAVQOVLQKGTTPASKLI 240
Db 203 RLDDFINVLSYDLHGSEKFTGHNSPLSLPEDPKSSA-----YAMNYWRKLGAPSEKLI 257
QY 241 LGMPTVGRSTFLASSSDTRVGAATSGTGPPTKEGMLAYEVCSS--WKGATKQRIQD 298
Db 258 MGITYGRTFRLKASKNGLOAIGPASPGRKYTKQAGFLAYFEICSFVW--GAKKHWDY 316
QY 299 OKVPYIFRDNQWGFDDVESFKTKVSLKQGLGAMVWALDLDPAFGSCNQGRYPLIQ 358
Db 317 QYVPYANKGEWGYDADISFSYKAWFIRREHFGAMVWTLDDVDRGFTCGTGPFLVY 376
QY 359 TL-----RQELSLPYLP-----SGTPELEVPKPGQSPSEHPGP 391
Db 377 VMNDILVRAEFSSTSLPQFWLSSAVNSSSTDPERLAVTKAWTTDIKILPG-----GEAGV 432
QY 392 SPQDFTFCQKADGLPNR 411
Db 433 -----TEHKGKCNMTITPR 447

RESULT 11
ID_OGP_HUMAN STANDARD; PRT; 678 AA.
AC Q12889; Q15841;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Oviduct-specific glycoprotein precursor (oviductal glycoprotein)
DE (Oviductin) (Estrogen-dependent oviduct protein).
GN OVGPI OR OGP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Oviduct;
RX MEDLINE=95119256; PubMed=7819450;
RA Arias E.B., Verhage H.G., Jaffe R.C.;
RT "Complementary deoxyribonucleic acid cloning and molecular
RT characterization of an estrogen-dependent human oviductal
RT glycoprotein."
RL Biol. Reprod. 51:685-694 (1994).
RN [2]
RP SEQUENCE FROM N.A.
RA Jaffe R.C.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: BINDS TO OOCYTE ZONA PELLUCIDA IN VIVO. MAY PLAY A ROLE
CC IN THE FERTILIZATION PROCESS AND/OR EARLY EMBRYONIC DEVELOPMENT.
CC -!- SUBCELLULAR LOCATION: Secretory granules.
CC -!- TISSUE SPECIFICITY: OVIDUCT.
CC -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; U09550; AAA86946.1; -
DR EMBL; U58010; AAB04126.1; -
DR EMBL; U58001; AAB04126.1; JOINED.
DR EMBL; U58002; AAB04126.1; JOINED.
DR EMBL; U58003; AAB04126.1; JOINED.

DR EMBL; U58004; AAB04126.1; JOINED.
DR EMBL; U58005; AAB04126.1; JOINED.
DR EMBL; U58006; AAB04126.1; JOINED.
DR EMBL; U58007; AAB04126.1; JOINED.
DR EMBL; U58008; AAB04126.1; JOINED.
DR EMBL; U58009; AAB04126.1; JOINED.
DR EMBL; U58010; AAB04126.1; JOINED.
DR MIM; HGNC:8524; OVGPI.
DR InterPro; IPR001579; Chitinase 18/2.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR PROSITE; PS01095; CHITINASE 18; FALSE_NEG.
KW Glycoprotein; Fertilization; Signal.
FT SIGNAL 1 21
FT CHAIN 22 678 OVIDUCT-SPECIFIC GLYCOPROTEIN.
FT CARBOHYD 402 402 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 580 580 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 596 596 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 648 648 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 477 477 M -> T (IN REF. 2).
FT CONFLICT 511 511 S -> P (IN REF. 2).
FT CONFLICT 514 514 Y -> H (IN REF. 2).
FT CONFLICT 676 676 E -> Q (IN REF. 2).
SQ SEQUENCE 678 AA; 75421 MW; 245F2CEDCE92768B CRC64;
Query Match 38.3%; Score 917.5; DB 1; Length 678;
Best Local Similarity 44.8%; Pred. No. 4.4e-62;
Matches 193; Conservative 66; Mismatches 145; Indels 27; Gaps 7;
QY 2 KLVCYFTNWAQVROGEARFLPKDLDPISLCTHLIYAFAGMTNHLSTTEWDE--TLYQEFN 60
Db 23 KLVCYFTNWAHSPGPASTLPHDLPFLCTHLIFAFASNNQIIVAKDLQDEKILYPERN 82
QY 61 GLKQWNPKLTLAIGGNWFGTKFTDMVATANNQTFVNSAIRFLKYSPFGLDLDWEY 120
Db 83 KLKERNRELKTLISIGGNFGTSRFTTLMSTFANREKFIASVLSLRTHDFDGLDLFFLY 142
QY 121 PGSGGSPAVDKERFTTLVODLANAFOEAQTSKGERLLLSAAVPAGQTYVDAGYVDKIA 180
Db 143 PGLRGSPMDHRTWFLFLIEELLFAFRKEALLTWPRLLLSAAVGVPHIVQTSYDVRFLG 202
QY 181 QNLDFVNLMAYPFHGSEKFTGHNSPLYKQESGAAASLVNDAVQOVLQKGTTPASKLI 240
Db 203 RLDDFINVLSYDLHGSEKFTGHNSPLSLPEDPKSSA-----YAMNYWRKLGAPSEKLI 257
QY 241 LGMPTVGRSTFLASSSDTRVGAATSGTGPPTKEGMLAYEVCSS--WKGATKQRIQD 298
Db 258 MGITYGRTFRLKASKNGLOAIGPASPGRKYTKQAGFLAYFEICSFVW--GAKKHWDY 316
QY 299 OKVPYIFRDNQWGFDDVESFKTKVSLKQGLGAMVWALDLDPAFGSCNQGRYPLIQ 358
Db 317 QYVPYANKGEWGYDADISFSYKAWFIRREHFGAMVWTLDDVDRGFTCGTGPFLVY 376
QY 359 TL-----RQELSLPYLP-----SGTPELEVPKPGQSPSEHPGPSPQD--TFQC 400
Db 377 VLNDILVRAEFSSTSLPQFWLSSAVNSSSTDPERLAVTKAWTTDTSKILPPGGEAGVTEH 436
QY 401 GKADGLYNPR 411
Db 437 GKCNMTITPR 447

RESULT 12
ID_CHIT_BRUMA STANDARD; PRT; 504 AA.
AC P29030;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE Endochitinase precursor (EC 3.2.1.14) (MFI antigen).
OS Brugia malayi.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
OX Onchocercidae; Brugia.
RN [1] NCBI_TaxID=6279;
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=92179220; PubMed=1542646;
RA Fuhrman J.A., Lane W.S., Smith R.F., Piesens W.F., Perler F.B.;
RT "Transmission-blocking antibodies recognize microfilarial chitinase
in brugian lymphatic filariasis";
RL Proc. Natl. Acad. Sci. U.S.A. 89:1548-1552(1992).
CC -!- FUNCTION: THE MFI ANTIGEN IS A MICROFILARIAL CHITINASE, WHICH MAY
FUNCTION TO DEGRADE CHITIN-CONTAINING STRUCTURES IN THE MICRO-
FILARIA OR IN ITS MOSQUITO VECTOR DURING PARASITE DEVELOPMENT AND
TRANSMISSION.
CC -!- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages of N-
acetyl-D-glucosamine polymers of chitin.
CC -!- DEVELOPMENTAL STAGE: THE APPEARANCE OF THE MFI ANTIGEN CORRESPOND
WITH THE ONSET OF THE PARASITE'S ABILITY TO INFECT THE MOSQUITO.
CC -!- PTM: O-GLYCOSYLATED.
CC -!- MISCELLANEOUS: KNOWN TO BIND CALCIUM.
CC -!- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
HYDROLASES).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; M73689; AAA27854.1; -;
DR PIR; A38221; A38221.
DR InterPro; IPR002557; Chitin_bind_Pera.
DR InterPro; IPR001579; Chitinase_18/2.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF01607; CBM_14; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR SMART; SM00494; ChtBD2; 1.
DR PROSITE; PS01095; CHITINASE_18; 1.
KW Hydrolyase; Glycosidase; Chitin degradation; Signal; Antigen;
KW Repeat; Glycoprotein; Calcium-binding.
FT SIGNAL 1 22
FT CHAIN 23 504 ENDOCHITINASE.
FT DOMAIN 23 400 CATALYTIC.
FT DOMAIN 401 450 SER/THR-RICH (LINKER).
FT DOMAIN 407 448 3 X 14 AA APPROXIMATE TANDEM REPEATS.
FT ACT SITE 148 / 148 PROTON DONOR (BY SIMILARITY).
SQ SEQUENCE 504 AA; 55971 MW; A78BE7BF8B9E3709B CRC64;
Query Match 35.9%; Score 860.5; DB 1; Length 504;
Best Local Similarity 37.0%; Pred. Mismatches 58;
Matches 177; Conservative 80; Nucleotide 170; Indels 51; Gaps 9;
5 CYFTNWAQVQGEARFLPKDPLSLCTHLIYAFAGMTHQLSTT---EWNDST-----L 55
27 CYFTNWAQVQGEARFLPKDPLSLCTHLIYAFAGMTHQLSTT---DELGDSKPFENWEDSTWSKM 84
56 YQEFNGLKQNPPLKTLTLLAIGGNFGTQKTDWVATANNRQTFVNSAIRFLKYSFDGLD 115
85 YSAVTKLRNTPGLKVLISYGGYNGFSAFTGTIAKSAQKTERFIKSAFLAKLNDFDGD 144
116 LDWEYEGSGSFAVDKERTTTLVQDLANAFQEAQTSKGERILLNAAVPAQTYVDAGYE 175
145 LDWEYFVGVA-----EEHAKLVEMAKMTAFVEAKTSKQRLTLTAASVAGKGTIDGSYN 198
176 VDKIAQNLDVNLMAVDYFHGSKEKVTGHSPLYKQESGSAASLNVDAAVQVQLQKGP 235
199 VESLGNFDLLFMSYDLHGSKEKVDLHGLKHPKTEGVSIGIFNTEPAADYNAKGM 258
236 ASKILGMPTYGRSTFLASSSDTRVCGAPATGSGTFCPTKEGMLAYVEVCSW-KGATKQ 294

Db 259 KEKIIIGIPYAQGWTLNPSETAIAAASRPSSASKTNPAGGTASYWEICKYLKEGGKE 318
Qy 295 RIQDKV-PYIFRDNQWGFDDVESFKTKVSLYKOKGLGGAMVWALDLDLDDPAGESCNOGR 353
Db 319 TVHQGVGAYWKGQWGYDNEETIRIKMKWKEGYGGAFIWDLDFDFTFGKSCGKGP 378
Qy 354 YPLIQLTRQELSLPYLPSTGTPLEVPKPG-----QPSE 386
Db 379 YPLLNAISSELE---GESENPEITTEPSITETAEVETDEETSETAEYDTDEETSE 435
Qy 387 PEHGPSQDFTFC--GXADGLYPNRERSFYSACARLFOOSCTGLVFSNSCKCC 442
Db 436 TEATTYDTDETEGQPCPERDGLFPHPTDCHLFIQANNIAYVMQCPATFFNDAIKVC 493
RESULT 13
CHIT MANSE
ID CHIT MANSE STANDARD; PRT; 554 AA.
AC P36362;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE Endochitinase precursor (EC 3.2.1.14).
OS Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata;
OC Ditrysia; Sphingioidea; Sphingidae; Sphinginae; Manduca.
OX NCBI_TaxID=7130;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93357793; PubMed=8353525;
RA Kramer K.J., Corpuz L., Choi H.K., Muthukrishnan S.;
RT "Sequence of a cDNA and expression of the gene encoding epidermal and
gut chitinases of Manduca sexta";
RL Insect Biochem. Mol. Biol. 23:691-701(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97215580; PubMed=9061927;
RA Choi H.K., Choi K.H., Kramer K.J., Muthukrishnan S.;
RT "Isolation and characterization of a genomic clone for the gene of an
insect molting enzyme, chitinase";
RL Insect Biochem. Mol. Biol. 27:37-47(1997).
CC -!- FUNCTION: DIGEST CHITIN IN THE EXOSKELETON DURING THE MOLTING
PROCESS.
CC -!- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages of N-
acetyl-D-glucosamine polymers of chitin.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: EPIDERMIS AND GUT.
CC -!- DEVELOPMENTAL STAGE: HIGH LEVELS SEEN IN THE EPIDERMIS ON DAY 0,
BUT RAPIDLY DISAPPEARS AND IS UNDETECTED ON DAYS 1-4 OF FIFTH
INSTAR. IT REAPPEARS ON DAY 5 AND PEAKS ON DAY 7 AFTER WHICH A
RAPID DECLINE IS SEEN. IN THE GUT IS DETECTED ON DAY 6 WITH LOWER
LEVELS SEEN ON DAYS 0, 7 AND 8.
CC -!- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
HYDROLASES).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; U02270; AAC04924.1; -;
DR EMBL; L49234; AAB53952.1; -;
DR InterPro; IPR002557; Chitin_bind_Pera.
DR InterPro; IPR001579; Chitinase_18/2.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF01607; CBM_14; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR SMART; SM00494; ChtBD2; 1.


```
DR PROSITE; PS01095; CHITINASE 18; 1.
KW Hydrolase; Glycosidase; Chitin degradation; Signal; Glycoprotein.
FT SIGNAL 1 39
FT CHAIN 20 554
FT DOMAIN 396 453
FT ACT SITE 146 146
FT CARBOHYD 85 85
FT CARBOHYD 303 303
FT CARBOHYD 407 407
FT CARBOHYD 545 545
SQ SEQUENCE 554 AA; 62203 MW; 3989D756C96CD490 CRC64;

Query Match 34.4%; Score 824; DB 1; Length 554;
Best Local Similarity 33.7%; Pred. No. 4.2e-55;
Matches 181; Conservative 87; Mismatches 159; Indels 110; Gaps 15;

QY 1 AKLVCFYTNWAOYRQGEARFLPKDLPSCSLTHLIYAFAGWTHQLSTTWNDETLL----- 55
DB 23 ARIVCFYSNMAVYRPGVGRYGEDIPVEKCTHIYSFIGV-----TEGNSVLLIDPE 75
QY 56 -----YQFENGLKQWPKLTLAIGGNWCTQKFTDMVATANNROTQVNSAIRFLRK 108
DB 76 LDVDKNGFRNFTSLRSSHPSVKPMVAVGWAEGSSKSHMAQKSTRMFSIRVSVSFLKK 135
QY 109 YSPDGLDLWEYPGS--QGSPAVDKERFTTLVODLANAQGEAQTSGKERLLLSAAVPAG 166
DB 136 YPFDGLDLWEYFGAADRGGSDKDKFLYLVELRRAFRVCK--GWE---LTAAPLA 190
QY 167 QTVYDAGYEVDKIAQNLDFVNLMAFYFHGSWEKVTGHNISPLYKQRESGAASLNVDAAV 226
DB 191 NFRLEMGYHVPCLQELDAIHVMSYDLRGNWAGFADVHSPLYKRPDQWAYEKLNVNDGL 250
QY 227 QWLQGTPTASKLILGMPYGRSFTLASSD-----TRVCAPATGSGTGPCTKEGMLA 281
DB 251 HLMEEKGSPSNKLWGPYFGRSFTLSAGNNVYGLGTFINKEA--GGDGPAPYNTATGFWA 309
QY 282 YVEVCSW-----KGATKQRIQDKQVPIFRDNQWGFDDVESKTVSYLKQKGLGAMV 336
DB 310 YVEICTEVDKDDSGWKWKDEQCKPYAYKGTQWGVYEDPRSVIEIKWNIKQKYLGMAT 369
QY 337 WALDLDPDFAGFCNQGRYPLIOTLRQLSLYLP-----SCTPELVKPKQPSPEH 389
DB 370 WAIMDDDFQGL--CGE--KNPLIKILHKHMSYVFPPTHTNTTTPPEWARP--PSTPSDPSE 426
QY 390 G-----PSPQDT--- 397
DB 427 GDIPTTTAKPASTTKVTKTTTTTAKPQSVDEENDINVRPEPKPEPEPEVEVP 486
QY 398 -----FCQKADGLYPNPRSSFSYSCAAGRLFOQSCPTGLVFSNCKCCTW 444
DB 487 PTENEVDGSEICNSDDYI--PDKKHCDKYRCVNGEAMQFSCQHGTVFVNLNVCDM 542

RESULT 14
CH11_BACCI
ID CH11_BACCI STANDARD; PRT; 699 AA.
AC P2053;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE Chitinase A1 precursor (EC 3.2.1.14).
GN CH11.
OS Bacillus circulans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1397;
RN [1]
RC SEQUENCE FROM N.A.
RX MEDLINE=90368776; PubMed=2203782;
RA Watanabe T., Suzuki K., Oyanagi W., Ohnishi K., Tanaka H.;
RT "Gene cloning of chitinase A1 from Bacillus circulans WL-12 revealed
its evolutionary relationship to Serratia chitinase and to the type
III homology units of fibronectin.";
```

```
RL J. Biol. Chem. 265:15659-15665 (1990).
RN [2]
RP MUTAGENESIS.
RC STRAIN-WL-12;
RX MEDLINE=93366760; PubMed=8103047;
RA Watanabe T., Kohori K., Miyashita K., Fujii T., Sakai H.,
RA Uchida M., Tanaka H.;
RT "Identification of glutamic acid 204 and aspartic acid 200 in
chitinase A1 of Bacillus circulans WL-12 as essential residues for
chitinase activity.";
RL J. Biol. Chem. 268:18567-18572 (1993).
CC -!- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages of N-
acetyl-D-glucosamine polymers of chitin.
CC -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC -!- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
HYDROLASES).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; M57601; AAA81528.1; .
DR PIR; A38368; A38368.
DR HSSP; P02751; LFNA.
DR InterPro; IPR003610; CEM_5_12.
DR InterPro; IPR001579; Chitinase_18/2.
DR InterPro; IPR003961; FN.III.
DR InterPro; IPR003962; FN.II repeat.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF00041; fn3; 2.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR Pfam; PF02839; CEM_5_12; 1.
DR PRINTS; PRO0014; FNTYPEI.II.
DR ProDom; PD000471; Glyco_hydro_18; 2.
DR SMART; SM00495; ChtBD3; 1.
DR SMART; SM00060; FN3; 2.
DR PROSITE; PS01095; CHITINASE 18; 1.
KW Hydrolase; Glycosidase; Chitin degradation; Signal; Repeat.
FT SIGNAL 1 41
FT CHAIN 42 699
FT DOMAIN 42 460
FT DOMAIN 465 549
FT DOMAIN 560 644
FT ACT SITE 204 204
FT MUTAGEN 200 200
FT MUTAGEN 204 204
FT SEQUENCE 699 AA; 73677 MW; AC7C9B2E2E987643 CRC64;

Query Match 24.4%; Score 584.5; DB 1; Length 699;
Best Local Similarity 33.0%; Pred. No. 9.3e-37;
Matches 147; Conservative 71; Mismatches 143; Indels 85; Gaps 16;

QY 2 KLVCYFTWAOYRQGEARFLPKDLPSCSLTHLIYAFAGW----- 40
DB 45 KIVGYFSAAYGR---NYNVADIDPTKVTHTNYAFADICWNGIHGNDPPSPNVTWTC 101
QY 41 TNHOLSTTE-----WMD-----ETLYQ-----EFNGLKQWPKLTLAI 75
DB 102 QNEKQTNVNPNGTIVLGDPMIDTGTAGTDWQPIAGNINQLNKLKQTNENLKIISV 161
QY 76 GGNFQTKFTDMVATANNROTQVNSAIRFLKYSFQDGLDLWEYPGSQ-----SPAVD 130
DB 162 GGMTW--SNRFSVAAATAATREVFANSAVDFLKYNFQDGLDLWEYFVSGGLDGNKRPED 220
QY 131 KERFTTLVODLANAQGEAQTSGKERLLLSAAVPAGQTVYDAGYEVDKIAQNLDFVNLMA 190
DB 221 KONYTILLISKIREKLDAAGAVDQK--KYLTIASGASATYA--ANTELAKIAIVDWINIWT 278
```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2003, 20:55:24 ; Search time 25.0148 Seconds
(without alignments)
3665.470 Million cell updates/sec

Title: US-10-004-219b-10

Perfect score: 2398

Sequence: 1 AKLVCFYTNWQYRQGEARF.....QQSCPTGLVFSNCKRCTWN 445

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_virius.*
- 16: sp_bacteriaph.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2398	100.0	466	4	Q13231	Q13231 homo sapien
2	1935	80.7	387	4	Q9H3V8	Q9H3V8 homo sapien
3	1334	55.6	396	11	Q9D7Q1	Q9D7Q1 mus musculus
4	1323	55.2	476	4	Q9B2P6	Q9B2P6 homo sapien
5	1309.5	54.6	472	11	Q9JLN1	Q9JLN1 mus musculus
6	1309.5	54.6	473	11	Q9D803	Q9D803 mus musculus
7	1304	54.4	472	6	Q95M17	Q95M17 bos taurus
8	1302.5	54.3	488	13	Q90W34	Q90W34 bufo japoni
9	1301.5	54.3	473	11	Q99PH2	Q99PH2 mus musculus
10	1053.5	43.9	352	11	Q9WTV1	Q9WTV1 rattus norv
11	1046.5	43.6	383	4	Q96H17	Q96H17 homo sapien
12	1041.5	43.4	390	4	Q96F97	Q96F97 homo sapien
13	1029.5	42.9	383	6	Q8SFQ0	Q8SFQ0 capra hircu
14	1021.5	42.6	383	6	Q29411	Q29411 sus scrofa
15	1015.5	42.3	696	5	Q9V2V2	Q9V2V2 drosophila
16	1015.5	42.3	1013	5	Q960M0	Q960M0 drosophila

17	991.5	41.3	381	11	Q99J84	Q99J84 mus musculus
18	991	41.3	368	4	Q9ULY4	Q9ULY4 homo sapien
19	979	40.8	332	6	O18949	O18949 bos taurus
20	975.5	40.7	365	11	Q91X29	Q91X29 mus musculus
21	970.5	40.5	398	11	Q91298	Q91298 mus musculus
22	970	40.5	402	11	Q8VHG1	Q8VHG1 mus musculus
23	968.5	40.4	398	11	P70201	P70201 mus musculus
24	967.5	40.3	398	11	O35744	O35744 mus musculus
25	965	40.2	402	11	O8VH43	O8VH43 mus musculus
26	926.5	38.6	624	6	O19118	O19118 macaca mula
27	919.5	38.3	4498	5	Q9W2Z3	Q9W2Z3 drosophila
28	883	36.8	525	5	O44079	O44079 anopheles g
29	868.5	36.2	504	5	Q9NGK8	Q9NGK8 wuchereria
30	864.5	36.1	457	6	Q95LB3	Q95LB3 oryctolagus
31	864.5	36.1	483	5	O23737	O23737 chelonus sp
32	858	35.8	520	5	O17100	O17100 acanthochei
33	856	35.7	497	5	O25615	O25615 onchocerca
34	855	35.7	524	5	O27919	O27919 acanthochei
35	848	35.4	460	5	Q9W2M7	Q9W2M7 drosophila
36	827.5	34.5	553	5	P91731	P91731 hyphantria
37	823	34.3	543	5	Q9GR93	Q9GR93 bombyx mori
38	822	34.3	566	5	O8WR52	O8WR52 bombyx mori
39	821	34.2	552	5	Q9GV44	Q9GV44 spodoptera
40	820	34.2	543	5	Q9GV05	Q9GV05 bombyx mori
41	820	34.2	565	5	P90710	P90710 bombyx mori
42	818	34.1	544	5	Q9GQC4	Q9GQC4 bombyx mori
43	816.5	34.0	467	5	O15993	O15993 penaeus jap
44	814	33.9	565	5	Q9GPG9	Q9GPG9 bombyx mand
45	810	33.8	460	5	Q8WS95	Q8WS95 glossina mo

ALIGNMENTS

RESULT 1

ID	Q13231	PRELIMINARY;	PRT;	466 AA.
AC	Q13231;			
DC	01-NOV-1996 (TREMBLrel. 01, Created)			
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)			
DE	01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE	Chitotriosidase precursor.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=96064695; PubMed=7592832;			
RA	Boot R.G., Renkema G.H., Strijland A., van Zonneveld A.J., Aerts J.M.;			
RT	"Cloning of a cDNA encoding chitotriosidase, a human chitinase			
RT	produced by macrophages.";			
RL	J. Biol. Chem. 270:26252-26256(1995).			
CC	-!- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL			
CC	HYDROLASES).			
DR	EMBL; U29615; AAC50246.1; -			
DR	InterPro; IPR001579; Chitinase_18/2.			
DR	InterPro; IPR002557; Chitin_bind_Pera.			
DR	InterPro; IPR001223; Glyco_hydro_18.			
DR	Pfam; PF01607; CBM_14; 1.			
DR	Pfam; PF00704; Glyco_hydro_18; 1.			
DR	ProDom; PD000471; Glyco_hydro_18; 1.			
DR	SMART; SM00494; ChitBD2; 1.			
DR	PROSITE; PS01095; CHITINASE_18; 1.			
KW	Glycosidase; Hydrolase; Signal.			
FT	SIGNAL 1 21			POTENTIAL.
FT	CHAIN 22 466			CHITOTRIOSIDASE.
SQ	SEQUENCE 466 AA; 51681 MW; B4312D1E885E386D CRC64;			

Query Match 100.0%; Score 2398; DB 4; Length 466;
Best Local Similarity 100.0%; Pred. No. 1.le-177;
Matches 445; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKLVCFYFTNWAQYRQGEARFLPKDLPDLPSLCTHLIIYAFAGTNNHQLSTTEWNETLYQEFN 60
 DB 22 AKLVCFYFTNWAQYRQGEARFLPKDLPDLPSLCTHLIIYAFAGTNNHQLSTTEWNETLYQEFN 81
 QY 61 GLKKNPKLKTLLAIGGNFQTKFTDMVATANNRQTFVNSAIRFLRKYSFQGLDLDWEY 120
 DB 82 GLKKNPKLKTLLAIGGNFQTKFTDMVATANNRQTFVNSAIRFLRKYSFQGLDLDWEY 141
 QY 121 PGSQSPAVDKERFTTLVQDLANAFQEQEATSGKERLLLSAAVPAGQTYVDAGYEVDKIA 180
 DB 142 PGSQSPAVDKERFTTLVQDLANAFQEQEATSGKERLLLSAAVPAGQTYVDAGYEVDKIA 201
 QY 181 QNLDFVNLMAFYDGHSGWEKVTGHSNPLKYRQESGAAASLNVDAAVQOVLQKGTTPASKLI 240
 DB 202 QNLDFVNLMAFYDGHSGWEKVTGHSNPLKYRQESGAAASLNVDAAVQOVLQKGTTPASKLI 261
 QY 241 LGMPTYGRSFTLASSSDTRVGAPATGSGTPGPTKEGGMALAYEVCWSKGGATKQRIQDOK 300
 DB 262 LGMPTYGRSFTLASSSDTRVGAPATGSGTPGPTKEGGMALAYEVCWSKGGATKQRIQDOK 321
 QY 301 VPYIFRDNQWGFDDVESFRTKVSYLKQKGLGGMVWALDLDLDFAGFSCNQGRIPIQTL 360
 DB 322 VPYIFRDNQWGFDDVESFRTKVSYLKQKGLGGMVWALDLDLDFAGFSCNQGRIPIQTL 381
 QY 361 ROELSPLYPGSGTPELEVPKPGOPSEPHGSPGQDTFCQKADGLYPNPRSSSFVSCA 420
 DB 382 ROELSPLYPGSGTPELEVPKPGOPSEPHGSPGQDTFCQKADGLYPNPRSSSFVSCA 441
 QY 421 AGRLFQOCSCTGLVFSNSCKCCTWN 445
 DB 442 AGRLFQOCSCTGLVFSNSCKCCTWN 466

RESULT 2

Q9H3V8 ID Q9H3V8 PRELIMINARY; PRT; 387 AA.
 AC Q9H3V8
 DT 01-MAR-2001 (T-EMBLrel. 16, Created)
 DT 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)
 DT 01-MAR-2002 (T-EMBLrel. 20, Last annotation update)
 DE Chitotriosidase precursor.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96064695; PubMed=7592832;
 RA Boot R.G., Renkema G.H., Strijland A., van Zonneveld A.J., Aerts J.M.;
 RT "Cloning of a cDNA encoding chitotriosidase, a human chitinase
 produced by macrophages".
 RL J. Biol. Chem. 270:26252-26256(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98421482; PubMed=9748235;
 RA Boot R.G., Renkema G.H., Verhoeck M., Strijland A., Blik J.,
 RA de Meuleneester T.M., Mannens M.M., Aerts J.M.;
 RT "The human chitotriosidase gene. Nature of inherited enzyme
 deficiency".
 RL J. Biol. Chem. 273:25680-25685(1998).
 CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
 HYDROLASES).
 CC ENBL; U62662; AAG10644.1; -.
 DR InterPro; IPR001579; Chitinase_18/2.
 DR InterPro; IPR001223; Glyco_hydro_18.
 DR Pfam; PF00704; Glyco_hydro_18; 1.
 DR ProDom; PD000471; Glyco_hydro_18; 1.
 DR PROSITE; PS01095; CHITINASE_18; 1.
 KW Glycosidase; Hydrolase; Signal.
 FT SIGNAL
 CHAIN 1 21
 SIGNAL 22 387
 CHAIN 387 AA; 43133 MW; 03A272B8BC5E0D71 CRC64;
 SQ SEQUENCE

Query Match 80.7%; Score 1935; DB 4; Length 387;
 Best Local Similarity 99.7%; Pred. No. 6.8e-142;
 Matches 364; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AKLVCFYFTNWAQYRQGEARFLPKDLPDLPSLCTHLIIYAFAGTNNHQLSTTEWNETLYQEFN 60
 DB 22 AKLVCFYFTNWAQYRQGEARFLPKDLPDLPSLCTHLIIYAFAGTNNHQLSTTEWNETLYQEFN 81
 QY 61 GLKKNPKLKTLLAIGGNFQTKFTDMVATANNRQTFVNSAIRFLRKYSFQGLDLDWEY 120
 DB 82 GLKKNPKLKTLLAIGGNFQTKFTDMVATANNRQTFVNSAIRFLRKYSFQGLDLDWEY 141
 QY 121 PGSQSPAVDKERFTTLVQDLANAFQEQEATSGKERLLLSAAVPAGQTYVDAGYEVDKIA 180
 DB 142 PGSQSPAVDKERFTTLVQDLANAFQEQEATSGKERLLLSAAVPAGQTYVDAGYEVDKIA 201
 QY 181 QNLDFVNLMAFYDGHSGWEKVTGHSNPLKYRQESGAAASLNVDAAVQOVLQKGTTPASKLI 240
 DB 202 QNLDFVNLMAFYDGHSGWEKVTGHSNPLKYRQESGAAASLNVDAAVQOVLQKGTTPASKLI 261
 QY 241 LGMPTYGRSFTLASSSDTRVGAPATGSGTPGPTKEGGMALAYEVCWSKGGATKQRIQDOK 300
 DB 262 LGMPTYGRSFTLASSSDTRVGAPATGSGTPGPTKEGGMALAYEVCWSKGGATKQRIQDOK 321
 QY 301 VPYIFRDNQWGFDDVESFRTKVSYLKQKGLGGMVWALDLDLDFAGFSCNQGRIPIQTL 360
 DB 322 VPYIFRDNQWGFDDVESFRTKVSYLKQKGLGGMVWALDLDLDFAGFSCNQGRIPIQTL 381
 QY 361 ROELS 365
 DB 382 ROELN 386
 RESULT 3
 Q9D7Q1 ID Q9D7Q1 PRELIMINARY; PRT; 396 AA.
 AC Q9D7Q1
 DT 01-JUN-2001 (T-EMBLrel. 17, Created)
 DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)
 DT 01-MAR-2002 (T-EMBLrel. 20, Last annotation update)
 DE 2300002L19RIK protein.
 GN 2300002L19RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=C57BL/6J; TISSUE=TONGUE;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staabli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 RA Hayaishizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
 HYDROLASES).
 CC ENBL; AK009012; BAB26025.1; -.

```
DR MGD; MGI:1919134; 2300002L19Rik.
DR InterPro; IPR000677; 2S Globulin.
DR InterPro; IPR001579; Chitinase 18/2.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR PRINTS; PR00551; 2SGLOBULIN.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR PROSITE; PS01095; CHITINASE_18; 1.
KW Glycosidase; Hydrolase.
SQ SEQUENCE 396 AA; 43413 MW; 32BB674B61AF7123 CRC64;

Query Match 55.6%; Score 1334; DB 11; Length 396;
Best Local Similarity 76.4%; Pred. No. 2.9e-95;
Matches 249; Conservative 26; Mismatches 49; Indels 2; Gaps 1;

QY 1 AKLVCFYTNWAOYRQGEARFLPKDLPSCLTLLIYAFAGMTHQSLSTTEWNETLYQEFN 60
DB 22 AKLVCFYTNWAOYRQGEARFLPKDLPSCLTLLIYAFAGMTHQSLSTTEWNETLYQEFN 81
QY 61 GLKXNPCLKTLLAIGGMNFGTQKFTDMVATANNRQTFVNSAIRFLRKYKSFQGLDLDMWEY 120
DB 82 SLKTNPKLTKLLAVGWTFTGQKFTDMVATASNRQTFVKSALSFLRTQGFQGLDLDMWEY 141
QY 121 PQSQSPAVDKERFTLVQDLANAFQGEAQTSGKERLLLSAAVPAQOTYVDAGYEVDKIA 180
DB 142 PQGRGSPVTKERFTALIQDLAKAFQGEAQTSGKERLLLSAAVPAQOTYVDAGYEVDKIA 201
QY 181 QNLDFNLWMAYPHGSWEKVTGNSPLYKROESGAASLNVDAAVQWLQKGTTPASKLI 240
DB 202 QSLDFNLWMAYPHGSWEKVTGNSPLYKROESGAASLNVDAAVQWLQKGTTPASKLI 261
QY 241 LGMPTYGRSFTLASSDTRVGAPATGSGTPGPTKEGGLMAYVEVCSW-KGATKQRIQDQK 300
DB 262 LGMPTYGRSFTLASSDTRVGAPATGSGTPGPTKEGGLMAYVEVCSW-KGATKQRIQDQK 319
QY 301 VPIYFRDNQWVGDDVESFKTKVSYL 326
DB 320 VPVAFQDNQWVGDDVESFKTKVSYL 345

RESULT 4
Q9BZP6 PRELIMINARY; PRT; 476 AA.
AC Q9BZP6;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE Putative chitinase precursor (Fragment).
DE Acidic mammalian chitinase precursor (SC 3.2.1.14).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=STOMACH, AND LUNG;
RX MEDLINE=21125893; PubMed=11085997;
RA Boot R.G., Blommaert E.F.C., Swart E., Ghaubharali-Van Der Vlugt K.,
RA Bijl N., Moe C., Place A., Aerts J.M.P.G.;
RT Identification of a Novel Acidic Mammalian Chitinase Distinct from
RT Chitotriosidase."
RL J. Biol. Chem. 276:6770-6778 (2001).
DR EMBL; AF290004; AAG60019.1; -.
DR InterPro; IPR001579; Chitinase 18/2.
DR InterPro; IPR002557; Chitin bind PerA.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF01607; CBM_14; 1.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR SMART; SM00494; ChtBD2_1.
DR PROSITE; PS01095; CHITINASE_18; UNKNOWN_1.
KW Glycosidase; Hydrolase; Signal.
FT SIGNAL 1 21 POTENTIAL.
SQ SEQUENCE 476 AA; 52271 MW; 92B27BAD2F7EB4CC CRC64;

Query Match 55.2%; Score 1323; DB 4; Length 476;
Best Local Similarity 52.2%; Pred. No. 2.8e-94;
Matches 240; Conservative 70; Mismatches 126; Indels 24; Gaps 3;

QY 2 KLVCFYTNWAOYRQGEARFLPKDLPSCLTLLIYAFAGMTHQSLSTTEWNETLYQEFN 61
DB 23 QLVCFYTNWAOYRQGEARFLPKDLPSCLTLLIYAFAGMTHQSLSTTEWNETLYQEFN 82
QY 62 LKXNPCLKTLLAIGGMNFGTQKFTDMVATANNRQTFVNSAIRFLRKYKSFQGLDLDMWEY 121
DB 83 LKXNSQKLTLLAIGGMNFGTQKFTDMVATANNRQTFVNSAIRFLRKYKSFQGLDLDMWEY 142
QY 122 SQSQSPAVDKERFTLVQDLANAFQGEAQTSGKERLLLSAAVPAQOTYVDAGYEVDKIA 181
DB 143 GSRGSPDPKHLFTLVQEMREARPEAKQINPKRLMVTAAVAAAGISNIQSGYEIPQLSQ 202
QY 182 NLDVNLWMAYPHGSWEKVTGNSPLYKROESGAASLNVDAAVQWLQKGTTPASKLI 241
DB 203 YLDYTHVMTYDLHGSWEKVTGNSPLYKRYPTDTGNSNLYLVNVDYVNNYKWDNGAPAEKLI 262
QY 242 GMPTYGRSFTLASSDTRVGAPATGSGTPGPTKEGGLMAYVEVCSW-KGATKQRIQDQ 299
DB 263 GPTYGRSFTLASSDTRVGAPATGSGTPGPTKEGGLMAYVEVCSW-KGATKQRIQDQ 322
QY 300 KVPYIFRDNQWVGDDVESFKTKVSYLQKGLGGMVWALDLDLPAGFCNQGRIPLIQT 359
DB 323 EVPYAYQVNWVGVDNIKSFDIKAOQLKXNKFGGAMWALDLDFTGTFCNKGKPLIST 382
QY 360 LRQELSLPYLPSGTPELEVPKQPSBP-----EHGSPSGQDTFCQKAD 404
DB 383 LKAL-----GLQASACTAPAQIEPTAAPSNGSGSSSSSGSGSGGFCVAVRAN 435
QY 405 GLYPNPRSSFSYSCAAGRLFOOSCTGLVSNCKCCTW 444
DB 436 GLYPVANNRNFHWCNGVTYQONCQAGLVFDTSCDCCNW 475

RESULT 5
Q9JLN1 PRELIMINARY; PRT; 472 AA.
AC Q9JLN1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Putative chitinase precursor (Fragment).
DE CHIA OR YNL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC1476668; TISSUE=SKIN;
RA Price P.A., Harris S.C., Williamson M.K.;
RT "YNL, A Putative Mouse Chitinase."
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
CC - SIMILARITY; BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
CC HYDROLASES).
DR EMBL; AF154571; AAF31644.1; -.
DR MGD; MGI:1932052; Chia.
DR InterPro; IPR001579; Chitinase 18/2.
DR InterPro; IPR002557; Chitin bind PerA.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF01607; CBM_14; 1.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR SMART; SM00494; ChtBD2_1.
DR PROSITE; PS01095; CHITINASE_18; 1.
KW Glycosidase; Hydrolase; Signal.
FT NON_TER 1 1
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 472 PUTATIVE CHITINASE.
```

```
SQ SEQUENCE 472 AA; 51872 MW; FF59088512C8A7F0 CRC64;
Query Match 54.6%; Score 1309.5; DB 11; Length 472;
Best Local Similarity 51.8%; Pred. No. 3.1e-93;
Matches 236; Conservative 71; Mismatches 128; Indels 21; Gaps 3;

QY 3 LVCYFTNWAQYRQGEARFLPKDLPDLCLTHLIYAFAGMTNHQSLSTWENDETLIQEFGNGL 62
DQ 23 LICYFTNWAQYRGLSGFEDDINPCLCTHLYAFAGMNNETITWENDVTLKAFNDL 82
QY 63 KQWPKLTKLLAIGWNGFTQKFTDMVATANNROTTFVNSAIRFLRKYSPDGLDLDWEYFG 122
DQ 83 KWRNKLTKLLAIGWNGFTQKFTDMVATANNROTTFVNSAIRFLRKYSPDGLDLDWEYFG 142
QY 123 SQGSPAVDKERFTTLVQDLANAFQOEAQTSKGERLLLSAAVPAQOTYVDAGYEVDKIAQN 182
DQ 143 SRGSPQDKHLFTLVKEMREAFQEAIESNRPRMLVMTAAVAGGINSIQAGYIEIPLSKY 202
QY 183 LDFVNLMAVDYFHGSWEKVTGHNPSPLKQESGAASLNVDAAVQOQLQGTTPASKLIILG 242
DQ 203 LDFHVMYDLHGSWEKVTGHNPSPLKQESGAASLNVDAAVQOQLQGTTPASKLIILG 262
QY 243 MPTYGRSFTLASSDTRVGPATGTPGFTKEGGMALYEVCSW--KGATKQRIQDOK 300
DQ 263 FPEYGHFTILRNPDSNGIGAPTSQDGPAGPYTRQAGFWAYEICTFLRSGATEVWDASOE 322
QY 301 VPIYFRDNQWVGFDDVESFKTKVSKYLKQKGLGAMVWALDLDLDFAGFSCNQGRIPLIOTL 360
DQ 323 VPYAYKANWLGVDNISKFSVKAQWLKQNNFGGAMIWALDLDLDFAGFSCNQGRIPLIOTL 382
QY 361 ROELSPLYPGSGTPELVKPGQPSBP-----EHPGSPGQDTFCQKADGLYP 408
DQ 383 NKALGISTEGCTAPDV-----PSEPVTPPGSGSGSGSGSGSGSGSGSGSGSGSGSGSG 435
QY 409 NPRESSFSVCAAGRLFOQSCPTGLVFSNSCKCTW 444
DQ 436 VADDRNFAWQINGITYQOHCQAGLVFDTSCNCCNW 471

RESULT 6
Q9D803 PRELIMINARY; PRT; 473 AA.
AC Q9D803;
DT 01-JUN-2001 (T-EMBLrel. 17, Created)
DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)
DE 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
DE 2200003E03RIK protein.
GN CHIA OR 2200003E03RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=STOMACH;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriam L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Bojelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Maehima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz K., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
```

```
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
CC HYDROLASES).
DR EMBL; AK008633; BAB25795.1; -.
DR MGD; MGI:1932052; Chia.
DR InterPro; IPR001579; Chitinase_18/2.
DR InterPro; IPR002557; Chitin_bind_Pera.
DR InterPro; IPR001223; Glyco_Hydro_18.
DR Pfam; PF01607; CBM_14; 1.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR SMART; SM00494; ChtBD2; 1.
DR PROSITE; PS01095; CHITINASE_18; 1.
KW Glycosidase; Hydrolase.
SQ SEQUENCE 473 AA; 52003 MW; 333C874477476695 CRC64;

Query Match 54.6%; Score 1309.5; DB 11; Length 473;
Best Local Similarity 51.8%; Pred. No. 3.1e-93;
Matches 236; Conservative 71; Mismatches 128; Indels 21; Gaps 3;

QY 3 LVCYFTNWAQYRQGEARFLPKDLPDLCLTHLIYAFAGMTNHQSLSTWENDETLIQEFGNGL 62
DQ 24 LICYFTNWAQYRGLSGFEDDINPCLCTHLYAFAGMNNETITWENDVTLKAFNDL 83
QY 63 KQWPKLTKLLAIGWNGFTQKFTDMVATANNROTTFVNSAIRFLRKYSPDGLDLDWEYFG 122
DQ 84 KWRNKLTKLLAIGWNGFTQKFTDMVATANNROTTFVNSAIRFLRKYSPDGLDLDWEYFG 143
QY 123 SQGSPAVDKERFTTLVQDLANAFQOEAQTSKGERLLLSAAVPAQOTYVDAGYEVDKIAQN 182
DQ 144 SRGSPQDKHLFTLVKEMREAFQEAIESNRPRMLVMTAAVAGGINSIQAGYIEIPLSKY 203
QY 183 LDFVNLMAVDYFHGSWEKVTGHNPSPLKQESGAASLNVDAAVQOQLQGTTPASKLIILG 242
DQ 204 LDFHVMYDLHGSWEKVTGHNPSPLKQESGAASLNVDAAVQOQLQGTTPASKLIILG 263
QY 243 MPTYGRSFTLASSDTRVGPATGTPGFTKEGGMALYEVCSW--KGATKQRIQDOK 300
DQ 264 FPEYGHFTILRNPDSNGIGAPTSQDGPAGPYTRQAGFWAYEICTFLRSGATEVWDASOE 323
QY 301 VPIYFRDNQWVGFDDVESFKTKVSKYLKQKGLGAMVWALDLDLDFAGFSCNQGRIPLIOTL 360
DQ 324 VPYAYKANWLGVDNISKFSVKAQWLKQNNFGGAMIWALDLDLDFAGFSCNQGRIPLIOTL 383
QY 361 ROELSPLYPGSGTPELVKPGQPSBP-----EHPGSPGQDTFCQKADGLYP 408
DQ 384 NKALGISTEGCTAPDV-----PSEPVTPPGSGSGSGSGSGSGSGSGSGSGSGSGSGSG 436
QY 409 NPRESSFSVCAAGRLFOQSCPTGLVFSNSCKCTW 444
DQ 437 VADDRNFAWQINGITYQOHCQAGLVFDTSCNCCNW 472

RESULT 7
Q95M17 PRELIMINARY; PRT; 472 AA.
ID Q95M17;
AC Q95M17;
DT 01-DEC-2001 (T-EMBLrel. 19, Created)
DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
DE 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
DE Chitin binding protein b04.
GN CBP B04.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=LIVER;
RC MEDLINE=21475601; PubMed=11591385;
```

RA Suzuki M., Morimatsu M., Yamashita T., Iwanaga T., Syuto B.;
 RT "A novel serum chitinase that is expressed in bovine liver.";
 RL FEBS Lett. 506:127-130(2001).
 DR EMBL: AB051629; BAB71805.1; --
 DR InterPro; IPR001579; Chitinase_18/2.
 DR InterPro; IPR002557; Chitin_bind_Pera.
 DR InterPro; IPR001223; Glyco_hydro_18.
 DR Pfam; PF01607; CBM_14; 1.
 DR ProDom; PD00704; Glyco_hydro_18; 1.
 DR PROSITE; PS01095; CHITINASE_18; UNKNOWN_1.
 DR PROSITE; PS01095; CHITINASE_18; UNKNOWN_1.
 DR SIGNAL 1 18
 FT CHAIN 19 488 TOAD PANCREATIC CHITINASE.
 SQ SEQUENCE 472 AA; 52129 MW; 74A600E8DA04B1E CRC64;

Query Match 54.4%; Score 1304; DB 6; Length 472;
 Best Local Similarity 52.2%; Pred. No. 8.2e-93;
 Matches 239; Conservative 72; Mismatches 123; Indels 24; Gaps 5;
 QY 2 KLVCFYFTWAQVQGEARFLPKDLPSCILTHIYAFAGMTNHQSTTEWDDTLTQEFNG 61
 DB 23 QLVCFYFTWAQVQGEARFLPKDLPSCILTHIYAFAGMTNHQSTTEWDDTLTQEFNG 82
 QY 62 LKKMNPKLKTLAIGGWNFGTKFTDMVATANNROTFFVNSAIRFLRKYSFSGDLDDWEYP 121
 DB 83 LKKMNPKLKTLAIGGWNFGTKFTDMVATANNROTFFVNSAIRFLRKYSFSGDLDDWEYP 142
 QY 122 GSQSPAVDKERFTTLVQDLANAFQEAOTSGKERLLLSAAVPAQOTYVDAGYEVDKIAQ 181
 DB 143 GFRGSPQDHLFTLVQETRAFQEAQTKNPRLLVTAAGVSIQAQYEPQLSQ 202
 QY 182 NLDVFNLMAYDPHGSWEKVTGNSPLRYKQESGAASLNVDAAVQOVLQKTPASKLIL 241
 DB 203 YLDFTHVMTYDPHGSWEKVTGNSPLRYKQESGAASLNVDAAVQOVLQKTPASKLIL 262
 QY 242 GMPTYGRSFTLASSSDTRVGAPATGSGTPGPGFTKEGMLAYYEVCSW--KGATKQRIQDO 299
 DB 263 GPPAYGHNFILRDSNNGIGAPTSAGPAGPYTRAGFWAYYEICAFKLDGATEAWDDSQ 322
 QY 300 KVPYIFRDNQWGFDDVESFVKVSKYLGKGLGGAMVWALDDDFAGFCSCNORYPLIQT 359
 DB 323 NVPYAYKTEWGVYDYNVNSFRKAKWLKNNFGAMVWALDDDFAGFCSCNORYPLIQT 382
 QY 360 LRQELSLPLVPGSTPELEVPKQSPSEPHGSPGQDT-----FCQKADGL 406
 DB 383 LKDALG--LKSATCNAST-----QSSEP--NSSPGNESGKNSSSSEGRGYCAGKADGL 433
 QY 407 YNPNSRSFYSCAGRLFOQSCPTGLVFSNCKCCTW 444
 DB 434 YPVAONRANFNCVNGITYKQNCITGLVFDTSCHCCNW 471

RESULT 8
 Q90W34 PRELIMINARY; PRT; 488 AA.
 AC Q90W34;
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Toad pancreatic chitinase (EC 3.2.1.14).
 GN TPCASE.
 OS Bufo japonicus (Japanese toad).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Bufo.
 OX NCBI_TaxID=8387;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PANCREAS;
 RA Oshima H.;
 RL Thesis (2001), Department of Department of Biology, Waseda University,
 RL Tokyo, Japan.
 DR EMBL; AJ345054; CAC87888.1; --
 DR InterPro; IPR002086; Aldehyde_dehydr.
 DR InterPro; IPR001579; Chitinase_18/2.
 DR InterPro; IPR002557; Chitin_bind_Pera.

DR InterPro; IPR001223; Glyco_hydro_18.
 DR Pfam; PF01607; CBM_14; 1.
 DR ProDom; PD00704; Glyco_hydro_18; 1.
 DR PROSITE; PS000471; ALDERHYDE_DEHYDR_CYS; UNKNOWN_1.
 DR PROSITE; PS00070; ALDERHYDE_DEHYDR_CYS; UNKNOWN_1.
 DR PROSITE; PS01095; CHITINASE_18; UNKNOWN_1.
 DR PROSITE; PS01095; CHITINASE_18; UNKNOWN_1.
 DR SIGNAL 1 18
 FT CHAIN 19 488 TOAD PANCREATIC CHITINASE.
 SQ SEQUENCE 488 AA; 54319 MW; E75E1AD3CB2B4919 CRC64;
 Query Match 54.3%; Score 1302.5; DB 13; Length 488;
 Best Local Similarity 51.0%; Pred. No. 1.1e-92;
 Matches 240; Conservative 74; Mismatches 126; Indels 31; Gaps 6;
 QY 2 KLVCFYFTWAQVQGEARFLPKDLPSCILTHIYAFAGMTNHQSTTEWDDTLTQEFNG 61
 DB 20 KLVCFYFTWAQVQGEARFLPKDLPSCILTHIYAFAGMTNHQSTTEWDDTLTQEFNG 79
 QY 62 LKKMNPKLKTLAIGGWNFGTKFTDMVATANNROTFFVNSAIRFLRKYSFSGDLDDWEYP 121
 DB 80 LKKMNPKLKTLAIGGWNFGTKFTDMVATANNROTFFVNSAIRFLRKYSFSGDLDDWEYP 139
 QY 122 GSQSPAVDKERFTTLVQDLANAFQEAOTSGKERLLLSAAVPAQOTYVDAGYEVDKIAQ 181
 DB 140 GSRGSPEDKQFTVLIOEMLDAFNEEARSSGLPRLLITAAVSAAGKTIDAGYEIAKIQ 199
 QY 182 NLDVFNLMAYDPHGSWEKVTGNSPLRYKQESGAASLNVDAAVQOVLQKTPASKLIL 241
 DB 200 LLDFTVMTYDPHGSWEKVTGNSPLRYKQESGAASLNVDAAVQOVLQKTPASKLIL 259
 QY 242 GMPTYGRSFTLASSSDTRVGAPATGSGTPGPGFTKEGMLAYYEVCSW--KGATKQRIQDO 300
 DB 260 GFTYGRFTFRPNPNMCDVGPVSGAGSAGPYTRAGFWAYYEICATLWSSGTVKWPIDQR 319
 QY 301 VPIYIFRDNQWGFDDVESFVKVSKYLGKGLGGAMVWALDDDFAGFCSCNORYPLIQT 360
 DB 320 VPIYACKSNEWGFDDVESFVKVSKYLGKGLGGAMVWALDDDFAGFCSCNORYPLIQT 379
 QY 361 RQEL-----SLPYLPSGT-----PELEVPKQSPSEPHGSPGQD 396
 DB 380 KSLLEGSTVNPGEICGGISIPSTPSATTTTTTTAKPDCTTPEP--PVTTPVPV--PVID 436
 QY 397 ---TPCQKADGLVNPNSRSFYSCAGRLFOQSCPTGLVFSNCKCCTW 444
 DB 437 VDPNFCVEKTDGLHVNPNTNKFYICANGRTYSMKCADGLVFOASCNCNW 487

RESULT 9
 Q99PH2 PRELIMINARY; PRT; 473 AA.
 AC Q99PH2;
 DT 01-JUN-2001 (Tremblrel. 17, Created)
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Acidic mammalian chitinase (EC 3.2.1.14).
 GN CHIA.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALE/C; TISSUE=LUNG;
 RX MEDLINE=21125893; PubMed=11085997;
 RA Boot R.G.; Blommaert E.F.C.; Swart E.; Chauharali-Van Der Vlugt K.,
 RA Bijl N.; Moe C.; Place A.; Aerts J.M.F.G.;
 RT "Identification of a Novel Acidic Mammalian Chitinase Distinct from
 RT Chitotriosidase";
 RL J. Biol. Chem. 276:6770-6778(2001).
 CC -!- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
 CC HYDROLASES).
 DR EMBL; AF290003; AAG60018.1; --


```
DR MGD; MG1:1932052; Chia.
DR InterPro; IPR001579; Chitinase 18/2.
DR InterPro; IPR002557; Chitin bind PerA.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF01607; CBM_14; 1.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR SMART; SM00494; ChtBD2; 1.
DR PROSITE; PS01095; CHITINASE_18; 1.
KW Glycosidase; Hydrolase.
SQ SEQUENCE 473 AA; 51977 MW; 389D87557BEC0784 CRC64;

Query Match 54.3%; Score 1301.5; DB 11; Length 473;
Best Local Similarity 51.5%; Pred. No. 1.3e-92;
Matches 235; Conservative 71; Mismatches 129; Indels 21; Gaps 3;

QY 3 LVCYFTNWAQYRQGEARFLPKDLPDLPSLCTHLIYAFAGMTHQSLTTEWNDFTLYQEFNG 62
DB 1 LVCYFTNWAQYRQGEARFLPKDLPDLPSLCTHLIYAFAGMTHQSLTTEWNDFTLYQEFNG 62
QY 24 LVCYFTNWAQYRQGEARFLPKDLPDLPSLCTHLIYAFAGMTHQSLTTEWNDFTLYQEFNG 83
DB 24 LVCYFTNWAQYRQGEARFLPKDLPDLPSLCTHLIYAFAGMTHQSLTTEWNDFTLYQEFNG 83
QY 63 KKNMPLKLTLLAIGGNWFGTKFTDMVATANNROTTFVNSAIRFLRKYSFDCGLDLDWEY 122
DB 63 KKNMPLKLTLLAIGGNWFGTKFTDMVATANNROTTFVNSAIRFLRKYSFDCGLDLDWEY 122
QY 84 KKNMPLKLTLLAIGGNWFGTKFTDMVATANNROTTFVNSAIRFLRKYSFDCGLDLDWEY 143
DB 84 KKNMPLKLTLLAIGGNWFGTKFTDMVATANNROTTFVNSAIRFLRKYSFDCGLDLDWEY 143
QY 123 SQGSPAVDKERFTTLVQDLANAFQOEAGTSGKERLLLSAAVPAQOTYVDAGYEVDKIAQ 182
DB 123 SQGSPAVDKERFTTLVQDLANAFQOEAGTSGKERLLLSAAVPAQOTYVDAGYEVDKIAQ 182
QY 144 SRGSPQDKHLFTVLKEMREAEFEQAIESNRPLAVTRAAGGIGNIQAQYIPELSKY 203
DB 144 SRGSPQDKHLFTVLKEMREAEFEQAIESNRPLAVTRAAGGIGNIQAQYIPELSKY 203
QY 183 LDFVNLMAVDHFGSWKVTGCHNSPLKROESGAASLNVDAAVQOQLKGTGPKASKLIL 242
DB 183 LDFVNLMAVDHFGSWKVTGCHNSPLKROESGAASLNVDAAVQOQLKGTGPKASKLIL 242
QY 204 LDFHVMYDLHGSWEGYTGNSPLKYPETETSNAYLVNVMYKNGKNGAPAEKLI 263
DB 204 LDFHVMYDLHGSWEGYTGNSPLKYPETETSNAYLVNVMYKNGKNGAPAEKLI 263
QY 243 MPTYGRSFTLASSSDTRVGAPATGSGTPGPTKEGGMALAYEVCWSN--KGATKQRIODOK 300
DB 243 MPTYGRSFTLASSSDTRVGAPATGSGTPGPTKEGGMALAYEVCWSN--KGATKQRIODOK 300
QY 264 FPEYGHFTILRNPSDNGIGAPTSGDGPAGATYQAQFWAYEICTFLRSATGEVWDASQE 323
DB 264 FPEYGHFTILRNPSDNGIGAPTSGDGPAGATYQAQFWAYEICTFLRSATGEVWDASQE 323
QY 301 VPYIFRDNQWGFDDVESFKTKVSYLKQKGLGAMVWALDLDPAFSCNQGRYPLIQT 360
DB 301 VPYIFRDNQWGFDDVESFKTKVSYLKQKGLGAMVWALDLDPAFSCNQGRYPLIQT 360
QY 324 VPYAYKANWLGVDNIKFSVKAQWLKQNNFGGAMIWALDLDFTGSCDQKGFPLTSTL 383
DB 324 VPYAYKANWLGVDNIKFSVKAQWLKQNNFGGAMIWALDLDFTGSCDQKGFPLTSTL 383
QY 361 RQELSLPYLPSGTPELEVPKGPQSE-----EHSFGPDQTFCCQKADGLYP 408
DB 361 RQELSLPYLPSGTPELEVPKGPQSE-----EHSFGPDQTFCCQKADGLYP 408
QY 384 NKALGISTEGCTAPDV-----PSEFVTPPGSGSGSGSGSGSGSGSGSGSGSGSGSG 472
DB 384 NKALGISTEGCTAPDV-----PSEFVTPPGSGSGSGSGSGSGSGSGSGSGSGSGSG 472
QY 409 NPRESRFSYCAAGRLFOQSCPTGLVFNSSCKCTW 444
DB 409 NPRESRFSYCAAGRLFOQSCPTGLVFNSSCKCTW 444
QY 437 VADRNAFWQCGINGITYQHQCAGLVFDTSNCN 472
DB 437 VADRNAFWQCGINGITYQHQCAGLVFDTSNCN 472

RESULT 10
Q9WTV1 PRELIMINARY; PRT; 352 AA.
AC Q9WTV1;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Glycoprotein-39 (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=101116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LEWIS;
RA Wendling U.; Boots A.M.H.; van Eden W.;
RT "Cloning of the rat homologue of Human Cartilage glycoprotein-39 a
RT potential autoantigen in arthritis."
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF062038; AAD22610.1; -
DR HSSP; P07254; 1CTN.
DR InterPro; IPR000677; 2S Globulin.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF00704; Glyco_hydro_18; 1.

DR PRINTS; PRO0551; 2SGLOBULIN.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR NON_TER 1
FT NON_TER 352
FT NON_TER 352
SQ SEQUENCE 352 AA; 39391 MW; CBDE991610AC936C CRC64;

Query Match 43.9%; Score 1053.5; DB 11; Length 352;
Best Local Similarity 55.5%; Pred. No. 1.4e-73;
Matches 198; Conservative 63; Mismatches 87; Indels 9; Gaps 5;

QY 2 KLVCFYFTNWAQYRQGEARFLPKDLPDLPSLCTHLIYAFAGMTHQSLTTEWNDFTLYQEFNG 61
DB 2 KLVCFYFTNWAQYRQGEARFLPKDLPDLPSLCTHLIYAFAGMTHQSLTTEWNDFTLYQEFNG 61
QY 62 LKKNMPLKLTLLAIGGNWFGTKFTDMVATANNROTTFVNSAIRFLRKYSFDCGLDLDWEY 121
DB 62 LKKNMPLKLTLLAIGGNWFGTKFTDMVATANNROTTFVNSAIRFLRKYSFDCGLDLDWEY 121
QY 62 LKTRNPNLKTLLSVGVNFGSGRFSRISVNAKSRTFVQSVAPFLRTYGFDCGLDLAWLP 121
DB 62 LKTRNPNLKTLLSVGVNFGSGRFSRISVNAKSRTFVQSVAPFLRTYGFDCGLDLAWLP 121
QY 122 GSQSPAVDKERFTTLVQDLANAFQOEAGTSGKERLLLSAAVPAQOTYVDAGYEVDKIAQ 181
DB 122 GSQSPAVDKERFTTLVQDLANAFQOEAGTSGKERLLLSAAVPAQOTYVDAGYEVDKIAQ 181
QY 122 GPK-----DKQHFTTLIKELKAEFTKEVQ--PQTEKLLLSAAVSAVSGKVTLDSDGYDVAQIAQ 175
DB 122 GPK-----DKQHFTTLIKELKAEFTKEVQ--PQTEKLLLSAAVSAVSGKVTLDSDGYDVAQIAQ 175
QY 182 NLDVNLMAVDHFGSWKVTGCHNSPLKROESGAASLNVDAAVQOQLKGTGPKASKLIL 241
DB 182 NLDVNLMAVDHFGSWKVTGCHNSPLKROESGAASLNVDAAVQOQLKGTGPKASKLIL 241
QY 176 HUDFNILMTYDFHGTWRHTTGHHSPLFRQDTPGDFRFSNVVDYGVGYMLRLGAPTNNKLYM 235
DB 176 HUDFNILMTYDFHGTWRHTTGHHSPLFRQDTPGDFRFSNVVDYGVGYMLRLGAPTNNKLYM 235
QY 242 GMPTYGRSFTLASSSDTRVGAPATGSGTPGPTKEGGMALAYEVCWSN--KGATKQRIODOK 300
DB 242 GMPTYGRSFTLASSSDTRVGAPATGSGTPGPTKEGGMALAYEVCWSN--KGATKQRIODOK 300
QY 236 GIPTGKGFSTLA--SENVQGAFIGSGGLPGRYTKKGLAYYEICDFLRCGAEVHRLGQ 294
DB 236 GIPTGKGFSTLA--SENVQGAFIGSGGLPGRYTKKGLAYYEICDFLRCGAEVHRLGQ 294
QY 301 VPYIFRDNQWGFDDVESFKTKVSYLKQKGLGAMVWALDLDPAFSCNQGRYPL 356
DB 301 VPYIFRDNQWGFDDVESFKTKVSYLKQKGLGAMVWALDLDPAFSCNQGRYPL 356
QY 295 VPFATKGNQWGVDDPESVKNKVKYLNKQLAGAMVWALDLDPRGSCFCHNVHFL 351
DB 295 VPFATKGNQWGVDDPESVKNKVKYLNKQLAGAMVWALDLDPRGSCFCHNVHFL 351

RESULT 11
Q96HI7 PRELIMINARY; PRT; 383 AA.
AC Q96HI7;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Similar to chitinase 3-like 1 (cartilage glycoprotein-39).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC008568; AAH08568.1; -
DR InterPro; IPR001223; Glyco_hydro_18.
DR InterPro; IPR003662; sub transporter.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; UNKNOWN 1.
SQ SEQUENCE 383 AA; 42625 MW; 6C3EF133BDC7C2D1 CRC64;

Query Match 43.6%; Score 1046.5; DB 4; Length 383;
Best Local Similarity 53.6%; Pred. No. 5.6e-73;
Matches 196; Conservative 66; Mismatches 95; Indels 9; Gaps 5;

QY 2 KLVCFYFTNWAQYRQGEARFLPKDLPDLPSLCTHLIYAFAGMTHQSLTTEWNDFTLYQEFNG 61
DB 2 KLVCFYFTNWAQYRQGEARFLPKDLPDLPSLCTHLIYAFAGMTHQSLTTEWNDFTLYQEFNG 61
QY 62 LKKNMPLKLTLLAIGGNWFGTKFTDMVATANNROTTFVNSAIRFLRKYSFDCGLDLDWEY 121
DB 62 LKKNMPLKLTLLAIGGNWFGTKFTDMVATANNROTTFVNSAIRFLRKYSFDCGLDLDWEY 121
QY 83 LKTRNPNLKTLLSVGVNFGSGRFSRISVNAKSRTFVQSVAPFLRTYGFDCGLDLAWLP 142
DB 83 LKTRNPNLKTLLSVGVNFGSGRFSRISVNAKSRTFVQSVAPFLRTYGFDCGLDLAWLP 142
QY 122 GSQSPAVDKERFTTLVQDLANAFQOEAGTSGKERLLLSAAVPAQOTYVDAGYEVDKIAQ 181
DB 122 GSQSPAVDKERFTTLVQDLANAFQOEAGTSGKERLLLSAAVPAQOTYVDAGYEVDKIAQ 181
```

Db 143 GRR-----DKQHFTTLIKEMKAEFTKEAQ-PGKQQLLLSAAALSAGKVTIDSDYDIKISQ 196
 QY 182 NLDVFNLMAYDFHGSWEKVTGNSPLKYQEESSGAASLNVDAAVOOMLOKQTPASKLIL 241
 Db 197 HLDFTSLTYDFHGAWRGTTGHHSPFLFRQEDASDPDRFSNTDYAVGYMLRLGAPASKLVM 256
 QY 242 GMPTYGRSFTLASSSDTRVGPATSGTGPPTKXGGMLAYYEVCWS-KGATKQRIQDQ 300
 Db 257 GIPTGRSFTLA-SSETGVGAPISGPIGRFTKEAGTLAYYEICDFLRGATVHRILGQ 315
 QY 301 VPIYFRDNQWGFDDVESFTKYSYLKQKGLGAMVWALDLDLDDFAGFSCNQ-RYPLIQ 359
 Db 316 VPYATKGNQWGVDDQESVKSVQVQLKQLAGAMVWALDLDLDDFQSGFCQDLRPLPLTNA 375
 QY 360 LRQELS 365
 Db 376 IXDALA 381

RESULT 12

Q96F97 PRELIMINARY; PRT; 390 AA.
 AC Q96F97;
 DT 01-DEC-2001 (TremBLrel. 19, Created)
 DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TremBLrel. 20, Last annotation update)
 DE Chitinase 3-like 2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BRAIN;
 RA Strausberg R.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC011460; AAH11460.1; -
 DR InterPro; IPR001223; Glyco_hydro.18.
 DR Pfam; PF00704; Glyco_hydro.18; 1.
 DR ProDom; PD000471; Glyco_hydro.18; 1.
 SQ SEQUENCE 390 AA; 43529 MW; 6801E4AF721AD63F CRC64;

Query Match 43.4%; Score 1041.5; DB 4; Length 390;
 Best Local Similarity 51.9%; Pred. No. 1.4e-72;
 Matches 190; Conservative 76; Mismatches 91; Indels 9; Gaps 4;
 QY 2 KLVCYFTNWAQYRQGEARFLPKDLPSCLTLLIYAFAGTNNHQLSTTEWNETLYQEFNG 61
 Db 28 KLVCYFTNWSQDRQPGKFTPENIDPFLCSHLIYGFASIEENKVIKDKSEVMLYQTINS 87
 QY 62 LKKNPKLKTLLAIGGNWFGTQKFTDMVATANNRQTFVNSAIRFLKYSFDCGLDLDWEYP 121
 Db 88 LKTKNPKLKLISIGLYLFGFHPMDSSSTRLEFINSILFLRNHNFGLDVSWIYP 147
 QY 122 GSQSPAVDKERFTTLVODLANAFOEAACTSGKERLLLSAAVPAQTYVDAGYVDKIAQ 181
 Db 148 DQK-----ENTHTVLIHLAEAFQKFTKSTYKERLLLTIVGSAGROMIDNSYQVEKLA 202
 QY 182 NLDVFNLMAYDFHGSWEK--VTGNSPLKYQEESSGAASLNVDAAVOOMLOKQTPASKL 239
 Db 203 DLDFTNLLSDFPHGSWEKPLITGNSPLSKQDRGPSSYINVEYAVGYIHKGPSEKV 262
 QY 240 ILGMPTYGRSFTLASSSDTRVGPATSGTGPPTKXGGMLAYYEVCWS-KGATKQRIQD 298
 Db 263 VMGIPTYGRSFTLA-SAETTVGAPASGPGAAGPITSSSGFLAYYEICQFLKAKITRLQD 321
 QY 299 QKVPYIFRDNQWGFDDVESFTKYSYLKQKGLGAMVWALDLDLDDFAGFSCNQRYPLIQ 358
 Db 322 QQVPAVKNQWGVDDVKSMEIKVQFLKNLNLGAMVWISIDNDFTTGKSCNQGPPLVQ 381
 QY 359 LRQEL 364
 Db 376 IXDALA 381

Db 382 AVGRSL 387
 RESULT 13
 Q8SPQ0 PRELIMINARY; PRT; 383 AA.
 AC Q8SPQ0;
 DT 01-JUN-2002 (TremBLrel. 21, Created)
 DT 01-JUN-2002 (TremBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
 DE BP40 precursor.
 OS Capra hircus (Goat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
 OC Bovidae; Caprinae; Capra.
 OX NCBI_TaxID=9925;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=MAMMARY GLAND;
 RA Paramasivam M., Saravanan K., Sharma S., Mohanty A.K., Yadav S.,
 RA Srinivasan A., Singh T.P.;
 RT "Goat mammary gland mRNA for a novel protein, BP40, complete coding
 region.";
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY081150; AAL87007.1; -
 KW Signal.
 FT SIGNAL.
 SQ SEQUENCE 383 AA; 42893 MW; 17655ED4BE4E9F5F CRC64;
 Query Match 42.9%; Score 1029.5; DB 6; Length 383;
 Best Local Similarity 51.9%; Pred. No. 1.2e-71;
 Matches 190; Conservative 69; Mismatches 98; Indels 9; Gaps 5;
 QY 2 KLVCYFTNWAQYRQGEARFLPKDLPSCLTLLIYAFAGTNNHQLSTTEWNETLYQEFNG 61
 Db 23 KLCYFTSWSQYREGDGCSPDAIDPFLCTHIIYSPANISNEIDTWEMNDVLYTLNT 82
 QY 62 LKKNPKLKTLLAIGGNWFGTQKFTDMVATANNRQTFVNSAIRFLKYSFDCGLDLDWEYP 121
 Db 83 LKNRPKLTLLSVGNWFGPGRFSKIAKTSQSRFTFKSVPPFLRTHGFGDGLDLAWLYP 142
 QY 122 GSQSPAVDKERFTTLVODLANAFOEAACTSGKERLLLSAAVPAQTYVDAGYVDKIAQ 181
 Db 143 GRR-----DKRHLTGLVKEMKAEFAEAQ-AGTERLLLSAAVSAACKIAIDRCYDIAQISR 196
 QY 182 NLDVFNLMAYDFHGSWEKVTGNSPLKYQEESSGAASLNVDAAVOOMLOKQTPASKLIL 241
 Db 197 HLDFTSLTYDFHGAWRQTVGHHSPFLFRQEDASDRFSNADYAVSYMLRLGAPANKLVM 256
 QY 242 GMPTYGRSFTLASSSDTRVGPATSGTGPPTKXGGMLAYYEVCWS-KGATKQRIQDQ 300
 Db 257 GIPTGRSFTLASSK-TDVGAPISGPIGRFTKEAGTLAYYEICDFLRGATVHRILGQ 315
 QY 301 VPIYFRDNQWGFDDVESFTKYSYLKQKGLGAMVWALDLDLDDFAGFSCNQ-RYPLIQ 359
 Db 316 VPYATKGNQWGVDDQESVKVKARYLKNRQLAGAMVWALDLDLDDPRGTFCGONLTFPLTSA 375
 QY 360 LRQELS 365
 Db 376 VKDVLA 381
 RESULT 14
 Q29411 PRELIMINARY; PRT; 383 AA.
 AC Q29411;
 DT 01-NOV-1996 (TremBLrel. 01, Created)
 DT 01-NOV-1996 (TremBLrel. 01, Last sequence update)
 DT 01-MAR-2002 (TremBLrel. 20, Last annotation update)
 DE 38 kDa heparin-binding glycoprotein.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

```

OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SMOOTH MUSCLE;
RX MEDLINE=95286589; PubMed=7768902;
RA Shackleton L.M., Mann D.M., Millis A.J.;
RT "Identification of a 38-kDa heparin-binding glycoprotein (gp38k) in
RT differentiating vascular smooth muscle cells as a member of a group of
RT proteins associated with tissue remodeling.";
RL J. Biol. Chem. 270:13076-13083(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=SMOOTH MUSCLE;
RA Shackleton L.M., Mann D.M., Millis A.J.T.;
RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U19900; AAA86482.1; -
DR EMBL; Z47803; CAA8764.1; -
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
KW Heparin-binding.
SQ SEQUENCE 383 AA; 42443 MW; 3D1039F49910BDC4 CRC64;

Query Match 42.6%; Score 1021.5; DB 6; Length 383;
Best Local Similarity 52.2%; Pred. No. 4.9e-71;
Matches 191; Conservative 67; Mismatches 99; Indels 9; Gaps 5;

QY 2 KLVCYFTNNAOYRQGEARFLPKDLPSLCTHLIYAFAGTNHOLSTENDETLTYQFNG 61
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 62 LKQNPKLKTLAIGWNFGTQKFTDMVATANNROTFFVNSAIRFLRKYSPFGDGLDWEYP 121
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 83 LKQNPENLKLTLSSVGGWNGFSQSFKIASNTQSRFTFKSVPPFLTHGDFGLDLAWISP 142
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 122 GSGGSPAVDKERTFTLVQDLANAFQEAQTSKGERLLLSAAVPAQGYVDAGVEVDKIAQ 181
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 143 GRR-----DKRHLLTLVKEMKAEFVREA-LPGTERLLLSGAVSAGKVAIDRGVDIAQISQ 196
QY 182 NLDFVNLMAYDFHGSWEKVTGHNSPLYKQESGAASLNVDAAVQVQKQKGPASKLLIL 241
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 197 HLDFTISLLTYDFHGAQRQTTHGHSPLFRGGQDASSDFNSADYAVSYLRLGAPANKLV 256
QY 242 GMPYIGRSFTLASSSDTRVGAPATSGTGPFTFKGGMLAYEVCVSW-KGATKQRIQDQK 300
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 257 GIFTGRSFTLASSK-TDVGAPASGFGIPGRFTKKGILAYEICDFLQGATVRRPLGQQ 315
QY 301 VPYIFRDQNWGPDVDESFTKVKSYLKKQGLGAMVWALDLDLDFAGFSCNQG-RYPLIQT 359
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 316 VPYATKGNQWGVDDQESVKNKAKYLSKRLAGAMVWTLDDLDFRGNFCQNLRPLTSA 375
QY 360 LRQELS 365
DB ::||:
QY 376 IKDVL 381

```

RESULT 15

```

Q9VZV2 ID Q9VZV2 PRELIMINARY; PRT; 696 AA.
AC Q9VZV2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE CG1869 protein.
GN CG1869.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY.

```

```

RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celnikier S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA De Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostel C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glaeser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwac C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laslo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003477; AAF47714.1; -.
DR FlyBase; FBgn0035398; CG1869.
DR InterPro; IPR001579; Chitinase_18/2.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF00704; Glyco_hydro_18; 2.
DR ProDom; PD000471; Glyco_hydro_18; 2.
DR PROSITE; PS01095; CHITINASE_18; UNKNOWN 1.
SQ SEQUENCE 696 AA; 77366 MW; 6877585837A19CF6 CRC64;

Query Match 42.3%; Score 1015.5; DB 5; Length 696;
Best Local Similarity 53.0%; Pred. No. 3.6e-70;
Matches 197; Conservative 52; Mismatches 110; Indels 13; Gaps 5;

QY 2 KLVCYFTNNAOYRQGEARFLPKDLPSLCTHLIYAFAGTNHOLSTENDET-----L 55
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 124 KIVCYFTNNSQYRVKIGFVPEDIPADLCTHIIFAFGDKNKLSSYESNDETKNVPLG 183
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 156 YQEFNGLKKNPKLTLAIGWNFGTQKFTDMVATANNROTFFVNSAIRFLRKYSPFGDGLD 115
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 184 YERMVTLKKNPKLTLAIGWNFGTQKFTDMVATANNROTFFVNSAIRFLRKYSPFGDGLD 243
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 116 LDWEYFGSGSPAVDKERTFTLVQDLANAFQEAQTSKGERLLLSAAVPAQGYVDAGVE 175
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 244 MDWEYFP--KGSD--DKKNFVLLKELREAFEAQELKKPRLLLSAAVPAVPGDNRGGYD 299
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 176 VDKIAQNLDVFNLMAYDFHGSWEKVTGHNSPLYKQESGAASLNVDAAVQVQKQKGP 235
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 300 VPAIASYLDLFNLMAYDFHGSWEKVTGHNAPYASTSEWPKQLSVNNAASLWVMGAP 359
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 236 ASKLILGMPYIGRSFTLASSSDTRVGAPATSGTGPFTFKGGMLAYEVCVSW--KGATK 293
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 360 KEKLVIGMPTYIGRSFTLANPDHGNAPASGGREGVYTKEGGFLAYEICEMLNGAVY 419
DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 294 QRIOQKVPIYIFRDQNWGPDVDESFTKVKSYLKKQGLGAMVWALDLDLDFAGFSCNQG- 352

```

Db 420 VDDENKVEYLVGDQWVGDDERAIRNKKHWIKSNGFGAMVWTIDMDDFKGEVCGNV 479
Qy 353 RYPLJOTLRQEL 364
Db 480 KYPLIGAMREEL 491

Search completed: June 29, 2003, 21:01:47
Job time : 26.0148 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2003, 20:50:19 ; Search time 30.2885 Seconds
(without alignments)
2001.720 Million cell updates/sec

Title: US-10-004-219B-14
Perfect score: 2506
Sequence: 1 YQLTCYFTNWAQYRPLGRF.....QQNCOAGLVFTSDCCNWA 455

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 3
Maximum DB seq length: 4000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A. Geneseq 101002:*

- 1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2506	100.0	455	22	AA62541 Disease treatment
2	2506	100.0	476	22	AA62543 Disease treatment
3	2018	80.5	368	23	ABB7580 Novel human protei
4	1415	56.5	377	22	AA62545 Disease treatment
5	1415	56.5	398	22	AA62544 Disease treatment
6	1323	52.8	466	18	AAW08584 Human 50 kDa chiti
7	1323	52.8	466	19	AAW40259 Human chitinase pr
8	1323	52.8	466	20	AAW42425 MO-218 clone of hu
9	1323	52.8	466	22	AAE00432 Human chitinase pr
10	1323	52.8	466	23	ABB76291 Human chitinase.

11	1317	52.6	466	19	AAW40260 Human chitinase pr
12	1317	52.6	466	20	AAW42426 MO-13B clone of hu
13	1317	52.6	466	22	AAE00433 Human chitinase pr
14	1317	52.6	466	23	ABB76292 Human chitinase.
15	1316	52.5	466	18	AAW31498 Human chitotriosid
16	1297	51.8	464	18	AAW31497 Human chitotriosid
17	1177	47.0	373	19	AAW40262 Human chitinase pr
18	1177	47.0	373	19	AAW40261 Human chitinase pr
19	1177	47.0	373	20	AAW42427 Clone of the C-ter
20	1177	47.0	373	20	AAW42428 Chitinase amino ac
21	1177	47.0	373	23	ABB76293 Human chitinase tr
22	1177	47.0	373	23	ABB76294 Human chitinase an
23	1173	46.8	387	18	AAW08585 Human 39 kDa chiti
24	1011.5	40.4	718	16	AAW73992 Murine oviduct spe
25	1008.5	40.2	696	22	ABB58615 Drosophila melanog
26	995	39.7	203	22	AAW92862 Human digestive ey
27	978.5	39.0	421	19	AAW53122 Amino acid sequenc
28	970.5	38.7	385	19	AAW53123 Amino acid sequenc
29	970.5	38.7	385	19	AAW47033 Human cartilage gp
30	970.5	38.7	416	19	AAW47034 Human cartilage gp
31	970.5	38.7	421	19	AAW53121 Amino acid sequenc
32	970.5	38.7	423	18	AAW35930 Human cartilage gl
33	963.5	38.4	383	18	AAW26751 Bovine whey protei
34	961.5	38.4	383	21	AAW03442 Gp8k protein sequ
35	961.5	38.4	668	16	AAW73993 Hamster oviduct sp
36	943.5	37.6	537	16	AAW73991 Bovine oviduct spe
37	890.5	35.5	4498	22	ABB58595 Drosophila melanog
38	863	34.4	554	18	AAW01824 Manduca sexta larv
39	863	34.4	554	21	AAW07183 Manduca sexta gut
40	853	34.0	460	22	ABB64366 Drosophila melanog
41	768.5	30.7	595	22	ABB71737 Polypeptide involv
42	760	30.3	321	9	AAW81342 D. pteronyssius 98
43	711.5	28.4	490	21	AAW52535 Der. pteronyssius 98
44	711.5	28.4	490	23	AAW96339 Der. pteronyssius 98
45	711.5	28.4	509	21	AAW52533 D. pteronyssius 98

ALIGNMENTS

RESULT 1
AA62541
ID AA62541 standard; protein; 455 AA.
XX
AC AA62541;
XX
DT 24-AUG-2001 (first entry)
XX
DE Disease treatment related protein SEQ ID NO: 1.
XX
KW Disease treatment; infection; chronic occlusive pulmonary disease;
KW bronchial asthma.
XX
OS Homo sapiens.
XX
PN WO200136633-A1.
XX
PD 25-MAY-2001.
XX
PF 14-NOV-2000; 2000WO-JP08015.
XX
PR 15-NOV-1999; 99JP-0324467.
XX
PA (TAKE) TAKEDA CHEM IND LTD.
XX
PI Nakanishi A, Morita S;
XX
DR WPI; 2001-397791/42.
XX
PT New proteins, peptides and DNA for treatment of bronchial asthma,
PT chronic occlusive lung disease and infectious disease -
XX
PS Claim 1; Page 97-99; 114pp; Japanese.

XX The present invention provides the sequence of a protein which can be
CC used in the treatment and prevention of infectious diseases. Inhibitors
CC of the protein can be used to treat bronchial asthma and chronic
CC occlusive pulmonary disease. The present sequence is a protein described
CC in the exemplification of the invention.
XX
SQ Sequence 455 AA;
Query Match 100.0%; Score 2506; DB 22; Length 455;
Best Local Similarity 100.0%; Pred. No. 1.3e-219;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 YQLTCYFTNWAQYRPGGLGRFMPDNIDPCLCTHLIYAFAGRONNEITTIEMNDVTLQAFN 60
DB 1 YQLTCYFTNWAQYRPGGLGRFMPDNIDPCLCTHLIYAFAGRONNEITTIEMNDVTLQAFN 60
QY 61 GLKNKNSQLKTLAIGGNWFGTAPFTAMVSTPENRQTFTSVIKFLRQYEFDFGLDFDWEY 120
DB 61 GLKNKNSQLKTLAIGGNWFGTAPFTAMVSTPENRQTFTSVIKFLRQYEFDFGLDFDWEY 120
QY 121 PSRSGSPQDKHLFTVLVQEMREAFQEAQINKPRLMVTAAVAAGISNIQSGYEIPQLS 180
DB 121 PSRSGSPQDKHLFTVLVQEMREAFQEAQINKPRLMVTAAVAAGISNIQSGYEIPQLS 180
QY 181 QYLDYIHVMYDYLHGSWEGYTGNSPLKYPTDTGNSNAYLVNDVYVMYWKDNGAPAEKLI 240
DB 181 QYLDYIHVMYDYLHGSWEGYTGNSPLKYPTDTGNSNAYLVNDVYVMYWKDNGAPAEKLI 240
QY 241 VGPFTYGHNFILSNPNTGIGAPTSAGPAGPYAKESGIWAYYEICTFLKNGATQGWDA 300
DB 241 VGPFTYGHNFILSNPNTGIGAPTSAGPAGPYAKESGIWAYYEICTFLKNGATQGWDA 300
QY 301 QYVPYAYQGNVWVGYDNIKSFDIKAQWLKHNKFGGAMVWAIDLDDTGTFCNKGKPEPLIS 360
DB 301 QYVPYAYQGNVWVGYDNIKSFDIKAQWLKHNKFGGAMVWAIDLDDTGTFCNKGKPEPLIS 360
QY 361 TLKALGLQASCTAPAQPIETPTAAPSNGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSG 420
DB 361 TLKALGLQASCTAPAQPIETPTAAPSNGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSG 420
QY 421 NNRNAFWHCVNGVTYQONCQAGLVFTSDCCCNWA 455
DB 421 NNRNAFWHCVNGVTYQONCQAGLVFTSDCCCNWA 455
RESULT 2
AAG62543
ID AAG62543 standard; protein; 476 AA.
XX AAG62543;
DT 24-AUG-2001 (first entry)
XX
DE Disease treatment related protein SEQ ID NO: 5.
XX
KW Disease treatment; infection; chronic occlusive pulmonary disease;
KW bronchial asthma.
XX
OS Homo sapiens.
XX
PN WO200136633-A1.
XX
PD 25-MAY-2001.
XX
PF 14-NOV-2000; 2000WO-JP08015.
XX
PR 15-NOV-1999; 99JP-0324467.
XX
XX (TAKE) TAKEDA CHEM IND LTD.
XX
PI Nakanishi A, Morita S;
XX

DR WPI; 2001-397791/42.
XX New proteins, peptides and DNA for treatment of bronchial asthma,
PT chronic occlusive lung disease and infectious disease
XX
PS Example 5; Page 101-103; 114pp; Japanese.
XX
CC The present invention provides the sequence of a protein which can be
CC used in the treatment and prevention of infectious diseases. Inhibitors
CC of the protein can be used to treat bronchial asthma and chronic
CC occlusive pulmonary disease. The present sequence is a protein described
CC in the exemplification of the invention.
XX
SQ Sequence 476 AA;
Query Match 100.0%; Score 2506; DB 22; Length 476;
Best Local Similarity 100.0%; Pred. No. 1.4e-219;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 YQLTCYFTNWAQYRPGGLGRFMPDNIDPCLCTHLIYAFAGRONNEITTIEMNDVTLQAFN 60
DB 22 YQLTCYFTNWAQYRPGGLGRFMPDNIDPCLCTHLIYAFAGRONNEITTIEMNDVTLQAFN 81
QY 61 GLKNKNSQLKTLAIGGNWFGTAPFTAMVSTPENRQTFTSVIKFLRQYEFDFGLDFDWEY 120
DB 82 GLKNKNSQLKTLAIGGNWFGTAPFTAMVSTPENRQTFTSVIKFLRQYEFDFGLDFDWEY 141
QY 121 PSRSGSPQDKHLFTVLVQEMREAFQEAQINKPRLMVTAAVAAGISNIQSGYEIPQLS 180
DB 142 PSRSGSPQDKHLFTVLVQEMREAFQEAQINKPRLMVTAAVAAGISNIQSGYEIPQLS 201
QY 181 QYLDYIHVMYDYLHGSWEGYTGNSPLKYPTDTGNSNAYLVNDVYVMYWKDNGAPAEKLI 240
DB 202 QYLDYIHVMYDYLHGSWEGYTGNSPLKYPTDTGNSNAYLVNDVYVMYWKDNGAPAEKLI 261
QY 241 VGPFTYGHNFILSNPNTGIGAPTSAGPAGPYAKESGIWAYYEICTFLKNGATQGWDA 300
DB 262 VGPFTYGHNFILSNPNTGIGAPTSAGPAGPYAKESGIWAYYEICTFLKNGATQGWDA 321
QY 301 QYVPYAYQGNVWVGYDNIKSFDIKAQWLKHNKFGGAMVWAIDLDDTGTFCNKGKPEPLIS 360
DB 322 QYVPYAYQGNVWVGYDNIKSFDIKAQWLKHNKFGGAMVWAIDLDDTGTFCNKGKPEPLIS 381
QY 361 TLKALGLQASCTAPAQPIETPTAAPSNGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSG 420
DB 382 TLKALGLQASCTAPAQPIETPTAAPSNGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSG 441
QY 421 NNRNAFWHCVNGVTYQONCQAGLVFTSDCCCNWA 455
DB 442 NNRNAFWHCVNGVTYQONCQAGLVFTSDCCCNWA 476
RESULT 3
ABB97580
ID ABB97580 standard; protein; 368 AA.
XX
AC ABB97580;
XX
DT 27-JUN-2002 (first entry)
XX
DE Novel human protein SEQ ID NO: 848.
XX
KW Human; antianemic; vulnerary; antiinflammatory; immunomodulator;
KW antinfertility; cerebroprotective; cytostatic; rheumatic; gene therapy;
KW neuroprotective; antiparkinsonian; protein therapy; EST;
KW expressed sequence tag.
XX
OS Homo sapiens.
XX
PN WO200222660-A2.
XX
PD 21-MAR-2002.
XX

PF 10-SEP-2001; 2001WO-US26015.
XX
PR 11-SEP-2000; 2000US-0659671.
XX
PA (HYSE-) HYSEQ INC.
XX
Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX
DR WPI; 2002-292408/33.
DR N-PSDB; ABN32766.
XX
An isolated polynucleotide for treating diseases associated with its
PT encoded polypeptide such as cancer and multiple sclerosis -
XX
Claim 20; SEQ ID NO 848; 509pp; English.
XX
The present invention provides the protein and coding sequences of 444
CC novel human proteins. These were isolated from expressed sequences tags
CC (ESTs). They can be used to stimulate cell growth, to regulate
CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
CC e.g. in burn treatment, to regulate the immune system e.g. to treat
CC multiple sclerosis, to regulate activin or inhibin e.g. to treat
CC infertility, to regulate haemostasis or thrombolysis e.g. to treat
CC stroke and cancer, to screen for drugs, to treat inflammatory conditions
CC e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.
CC Parkinson's disease. The present sequence is a protein of the invention.
XX
SQ Sequence 368 AA;
Query Match 80.5%; Score 2018; DB 23; Length 368;
Best Local Similarity 100.0%; Pred. No. 3e-175;
Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 88 MVSTPENRTFTSVIKFLRQYEFDFGLDPDWEYPSGRGSPPODKHLFTVLVQEMREAFQ 147
Db 1 MVSTPENRTFTSVIKFLRQYEFDFGLDPDWEYPSGRGSPPODKHLFTVLVQEMREAFQ 60
QY 148 EAKQINKPLMTAAVAAGISIQSGYEIPOLSOYLDYIHVMYDILHGSWEGYTGNSPL 207
Db 61 EAKQINKPLMTAAVAAGISIQSGYEIPOLSOYLDYIHVMYDILHGSWEGYTGNSPL 120
QY 208 YKYPTDTSNAYLVNDVYVNMVYKNDGAPAEKLI VGPPTYGHNFI LSNPNTGIGAPTSGA 267
Db 121 YKYPTDTSNAYLVNDVYVNMVYKNDGAPAEKLI VGPPTYGHNFI LSNPNTGIGAPTSGA 180
QY 268 GPAGPYAKESGIWAYYEICTFLKNGATQGWDAPOEVYPAYQGNVWVGYDNIKSFDIKAQW 327
Db 181 GPAGPYAKESGIWAYYEICTFLKNGATQGWDAPOEVYPAYQGNVWVGYDNIKSFDIKAQW 240
QY 328 LKHNFEGGAMVWAIDLDDFTGTCNKGKPLISTLKKALGLOSASCTAPAOPIEPITRAP 387
Db 241 LKHNFEGGAMVWAIDLDDFTGTCNKGKPLISTLKKALGLOSASCTAPAOPIEPITRAP 300
QY 388 SSGSGSGSGSGSGSGSGGFCFCAVRANGLYPVANNRNFAFWHCVNGVYTYQNCQAGLVFTT 447
Db 301 SSGSGSGSGSGSGSGGFCFCAVRANGLYPVANNRNFAFWHCVNGVYTYQNCQAGLVFTT 360
QY 448 SDCDCNWA 455
Db 361 SDCDCNWA 368
RESULT 4
AAG62545
ID AAG62545 standard; protein; 377 AA.
XX
AC AAG62545;
XX
DT 24-AUG-2001 (first entry)
XX
DE Disease treatment related protein SEQ ID NO: 18.
XX

Disease treatment; infection; chronic occlusive pulmonary disease;
KW bronchial asthma.
XX
OS Mus sp.
XX
PN WO200136633-A1.
XX
PD 25-MAY-2001.
XX
PF 14-NOV-2000; 2000WO-JP08015.
XX
PR 15-NOV-1999; 99JP-0324467.
XX
PA (TAKE) TAKEDA CHEM IND LTD.
XX
PI Nakanishi A, Morita S;
XX
DR WPI; 2001-397791/42.
XX
New proteins, peptides and DNA for treatment of bronchial asthma,
PT chronic occlusive lung disease and infectious disease -
XX
Claim 18; Page 110-112; 114pp; Japanese.
XX
The present invention provides the sequence of a protein which can be
CC used in the treatment and prevention of infectious diseases. Inhibitors
CC of the protein can be used to treat bronchial asthma and chronic
CC occlusive pulmonary disease. The present sequence is a protein described
CC in the exemplification of the invention.
XX
SQ Sequence 377 AA;
Query Match 56.5%; Score 1415; DB 22; Length 377;
Best Local Similarity 68.1%; Pred. No. 3e-120;
Matches 256; Conservative 48; Mismatches 72; Indels 0; Gaps 0;
QY 1 YOLTCYFNWAOYRPLGLRFPNDIDPCLCTHLIYAFAGRNNEITTIEMNDVTLYQAFN 60
Db 1 YOLTCYFNWAOYRPLGLRFPNDIDPCLCTHLIYAFAGRNNEITTIEMNDVTLYQAFN 60
QY 61 GLKXNKSOLKTLAIGGNFPTAPTAMVSTPENRTFTSVIKFLRQYEFDFGLDPDWEY 120
Db 61 GLKXNKSOLKTLAIGGNFPTAPTAMVSTPENRTFTSVIKFLRQYEFDFGLDPDWEY 120
QY 121 PGSRSGPPODKHLFTVLVQEMREAFQEAQKQINKPLMTAAVAAGISIQSGYEIPOLS 180
Db 121 PGSRSGPPODKHLFTVLVQEMREAFQEAQKQINKPLMTAAVAAGISIQSGYEIPOLS 180
QY 181 QYLDYIHVMYDILHGSWEGYTGNSPLYKYPTDTSNAYLVNDVYVNMVYKNDGAPAEKLI 240
Db 181 QYLDYIHVMYDILHGSWEGYTGNSPLYKYPTDTSNAYLVNDVYVNMVYKNDGAPAEKLI 240
QY 241 VGFPTYGHNFI LSNPNTGIGAPTSGAGPAGPYAKESGIWAYYEICTFLKNGATQGWDA 300
Db 241 VGFPTYGHNFI LSNPNTGIGAPTSGAGPAGPYAKESGIWAYYEICTFLKNGATQGWDA 300
QY 301 QSVVPYAGNVMVGYDNIKSFDIKAQWLKHNKFGAMVWAIDLDDFTGTCNKGKPLIS 360
Db 301 QSVVPYAGNVMVGYDNIKSFDIKAQWLKHNKFGAMVWAIDLDDFTGTCNKGKPLIS 360
QY 361 TLKKALGLOSASCTAP 376
Db 361 TLKKALGLOSASCTAP 376
RESULT 5
AAG62544
ID AAG62544 standard; protein; 398 AA.
XX
AC AAG62544;
XX
DT 24-AUG-2001 (first entry)
XX

QY 242 GPPTYCHNFIILSNPNTGIGAPTSGAGPAGPYAKESGIWAYVEICTFLKNGATQGDAPQ 301
DB 263 GMPYGRSFTLASSSDTRVGAPATGTPGPTKEGMLAYVEVCSW--KGATKQRIQDQ 320
QY 302 EUPYAYQGNVWVGVDNINIKSFDIKAKWLKHNKFGGAMWAIDLDDTGTFCNQGKPEPLIST 361
DB 321 KVPYIFRDQNWVGDFDVSFKTKVSYLKQKGLGGAMVWALDLDPAFSCNQGRYPLIQT 380
QY 362 LKKAL-----GLQSASCTAPAOPIERTITAAPSGSGSGSGSGSGSGGFCAVRAN 414
DB 381 LRQELSLPYLPSGTPLEVPKPGQSEPE-----EHGPSFGQDTFCQGRAD 425
QY 415 GLYPVANNRNFHWCNGVYTYQONCOAGLVDFTSDCCNW 454
DB 426 GLYPNPRSSFYSCAAGRLFQSCPTGLVFSNCKCCTW 465

RESULT 7
ID AAW40259
XX AAW40259 standard; Protein; 466 AA.
AC AAW40259;
XX
DT 15-JUN-1998 (first entry)
XX Human chitinase protein from clone MO-218.
DE Chitinase; human; fungal infection; immunogen; diagnosis; treatment;
KW Gaucher's disease; transgenic; detection; hybridisation; antifungal;
KW rheumatoid arthritis; overexpression; extracellular matrix.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..21
FT Protein /label= signal
FT /label= chitinase
FT /note= "from clone MO-218"
XX
PN WO9747752-A1.
XX
PD 18-DEC-1997.
XX
PF 16-JUN-1997; 97WO-US10460.
XX
PR 14-JUN-1996; 96US-0663618.
XX
PA (ICOS-) ICOS CORP.
XX
PI Gray PW;
XX
XX WPI; 1998-052316/05.
DR N-PSDB; AAV10435.
XX
PT Nucleic acids encoding human chitinase - useful as antifungal
XX agents, especially in combination with other antifungals
XX
PS Claim 1; Page 41-42; 63pp; English.
XX
CC This sequence represents a novel human chitinase isolated from clone
CC MO-218. Chitinases are useful for treating or preventing fungal infection
CC and as immunogens for generating antibodies which are used to purify,
CC detect and quantify chitinases, e.g. for diagnosis of Gaucher's disease.
CC The nucleic acid sequence of the chitinase is also useful as a probe to
CC identify and isolate genomic DNA encoding chitinases or similar proteins,
CC or cells expressing them or to generate transgenic ('knockout') rodents.
CC It can also be used in hybridisation assays and to detect genetic
CC alterations in the chitinase gene related to disease. Agents that inhibit
CC this protein may be useful in treatment of Gaucher's disease and
CC rheumatoid arthritis, where overexpression of the protein can damage
CC the extracellular matrix. Chitinase also improves the activity of other
CC antifungal agents and may allow a reduction in the dose of such agents,

CC and thus of their side effects.
XX
SQ Sequence 466 AA;
Query Match 52.8%; Score 1323; DB 19; Length 466;
Best Local Similarity 52.2%; Pred. NO. 1e-111;
Matches 240; Conservative 70; Mismatches 126; Indels 24; Gaps 3;
QY 2 QLTCTVETNWAQYRCLGRFMSNDIDPCLCITLIIYAFAGRONNEITTIENWVDVLYQAFNG 61
DB 23 KLVCTFTTNWAQYRQEARFLPKDLDPSCLTLLIYAFAGTWNHQLSTTEWNETLYQENG 82
QY 62 LKNKNSQLKTLILAIGGMNFGTAPFTAMVSTPENRQTFITTSVIKLRQYEFDGLDFDWEYP 121
DB 83 LKXNPKLTKTLILAIGGMNFGTKFTDMVATANNRQTFVNSAIRFLKYSFDGLDLWEYP 142
QY 122 GSRGSPPODKHLFTVLVQEMREAFQEAQKQINPKRLMYTAAVAAGISNIQSGYEIPQLSQ 181
DB 143 GSGSPAVDKERFTTLVODLANAFQEAQTSKGERLLLSAAVPAGQTYVDAGYVDKIAQ 202
QY 182 YLDYIHWMTYDLHGSWEGYTGENSEPLYKYPTDTSNAYLVNVDYVNNYWKDNGAPAEKLI 241
DB 203 NLDFVNLWYDFHGSWEKVTGHNPLYKQESGAAASLNVDAVQOVLQKGTPLASKIL 262
QY 242 GFTYGHNFILSNPNTGIGAPTSGAGPAGPYAKESGIWAYVEICTFLKNGATQGDAPQ 301
DB 263 GMPYGRSFTLASSSDTRVGAPATGTPGPTKEGMLAYVEVCSW--KGATKQRIQDQ 320
QY 302 EUPYAYQGNVWVGVDNINIKSFDIKAKWLKHNKFGGAMWAIDLDDTGTFCNQGKPEPLIST 361
DB 321 KVPYIFRDQNWVGDFDVSFKTKVSYLKQKGLGGAMVWALDLDPAFSCNQGRYPLIQT 380
QY 362 LKKAL-----GLQSASCTAPAOPIERTITAAPSGSGSGSGSGSGSGGFCAVRAN 414
DB 381 LRQELSLPYLPSGTPLEVPKPGQSEPE-----EHGPSFGQDTFCQGRAD 425
QY 415 GLYPVANNRNFHWCNGVYTYQONCOAGLVDFTSDCCNW 454
DB 426 GLYPNPRSSFYSCAAGRLFQSCPTGLVFSNCKCCTW 465

RESULT 8
ID AAY42425
XX AAY42425 standard; Protein; 466 AA.
AC AAY42425;
XX
DT 10-DEC-1999 (first entry)
XX
DE MO-218 clone of human Chitinase, amino acid sequence.
XX
KW chitin; fungal infection; immunocompromised; AIDS; chemotherapy;
KW organ transplant; parasite; chitin-binding; allele; vector;
KW truncated protein; chitin binding region.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..21
FT Protein /label= Signal_peptide
FT /label= Mature protein
FT /note= "Human Chitinase"
FT Peptide 391..445
FT /label= Chitin binding
FT /note= "Chitin binding region"
XX
PN WO9946390-A1.
XX
PD 16-SEP-1999.
XX
PF 12-MAR-1999; 99WO-US05343.
XX

PR 12-MAR-1998; 98US-0039198.
XX (ICOS-) ICOS CORP.
XX Gray PW, Tjoelker LW;
XX WPI; 1999-551417/46.
DR N-PSDB; AAZ21847.
XX Novel chitin-binding fragments of human chitinase used to treat fungal
PT infections in animals
XX
PS Example 1; Page 58-59; 83pp; English.
XX
CC This is the amino acid sequence of an allelic form of the human
CC chitinase enzyme, which is capable of degrading Chitin (a linear
CC homo polymer of beta-1,4-linked N-acetylglucosamine residues).
CC The sequence also includes a Chitin binding peptide region, which lacks
CC any chitinase activity, but which can be used to target anti-fungal
CC agents to chitin containing fungal cells.
CC Chitinase fragments can be used to screen for proteins or other
CC molecules that specifically bind to the chitin-binding domain of human
CC chitinase or that modulate its activity. These compounds are useful for
CC immunization, as well as for purifying chitinase, as well as for
CC detection and quantification of chitinase. Polynucleotide fragments of
CC the invention are useful as a source of probes and primers, and to
CC express the proteins recombinantly. The chitinase fragments, when
CC conjugated to antifungal compounds, are used to treat animals,
CC especially humans, infected with chitin-containing parasites such as
CC fungi. Fungal infection treated include candidiasis, aspergillosis,
CC coccidioidomycosis, blastomycosis, paracoccidioidomycosis,
CC mucormycosis, histoplasmosis, cryptococcosis, chromoblastomycosis,
CC sporotrichosis, and dermatophytoses.
CC Chitin can be degraded by the enzyme chitinase. Use of whole chitinase
CC protein for treating infections, especially fungal infections, is
CC problematic. In view of the increasing incidents of life-threatening
CC fungal infection in e.g. immunocompromised individuals, there exists a
CC need for identifying new compounds for treating fungal infection. The
CC chitin-binding fragments of the present invention provide this need.
XX
XX Sequence 466 AA;
Query Match 52.8%; Score 1323; DB 20; Length 466;
Best Local Similarity 52.2%; Pred. No. 1e-111;
Matches 240; Conservative 70; Mismatches 126; Indels 24; Gaps 3;
2 QLTCTYTNNAQYRPGLRGMPNDIDCLCTHLYAFAGRONNEITWNVLYQAFNG 61
23 KLVCTYTNNAQYRQGEARFLPKDLPDLCTHLYAFAGMTNHLSTWNETLYQEFNG 82
62 LKNNKSQLTLLAIGWNFGTAPFTMVSTPENRQTFTSVIKFLQYEPFDGLDPEWYEP 121
83 LKQNPFLKTLAIGWNFGTQFTDMVANNRQTFVNSAIRFLKYRFDGLDPEWYEP 142
122 GSRGSPQDKHLFTVLVOEMREAFQEAQKINPKRLMTAAVAAGISNTQSGYEIPQLSQ 181
143 GSGQSPAVDKERTFTVLQDLANAFQEAQTSGRERLLLSAAVPAGQTYVDAGVEVDKIAQ 202
182 YLDYIHVMYDLHSGNEGYTGENSEPLKYPTDTGNSAYLNVDVVMYKDNKGAPEKLIV 241
203 NLDVFNLMAYDFHGSWEKVTGHNPSPLYKRBESGAASLNVDVAQVQMLQKGTFPASKLIL 262
242 GFTYGHNFILSPNTGIGAPTSAGPAGPYAKESGIWAYEYICITFLKNGATQGWDAPO 301
263 GMYTGRSFTLASSDTRVGAPATGSGTGPFTFKEGMLAYEVCW--KGATKQRIQDQ 320
302 EYVYAYQGNVWGYDNKSPDIIKAQWLKHNKFGAMVWAILDDFTGTCNQCKFPLIST 361
321 KVPYIFRQNWGVDFDVSFPTKVSYLKQKGLGCMVWALDDDFAGFSCNQGRYPLIQT 380
362 LKAL-----GLQASCTAPAQPIEPIITAAPSGSGNGSGSGSGSGGSCGFCVAVRAN 414
381 LRQELSLPYLPSTPELEVPKQPQSEP-----EHGFSFGQDTFCQKAD 425

QY 415 GLYPVANNRFAFWHCVNGVTYQNCQAGLVFDTSCDCNNW 454
DB 426 GLYPNPRSSPYSCAAGRLFQSCPTGLVFSNCKCCTW 465

RESULT 9

AAE00432
ID AAE00432 standard; Protein; 466 AA.
XX
AC AAE00432;
XX
DT 19-JUN-2001 (first entry)
XX
DE Human chitinase protein from clone pMO-218.
XX
DE Human; antifungal; chitinase; immunoglobulin; Ig; therapy;
KW fungal infection; candidiasis; aspergillosis; coccidioidomycosis;
KW blastomycosis; paracoccidioidomycosis; histoplasmosis; cryptococcosis;
KW chromoblastomycosis; sporotrichosis; mucormycosis; dermatophytoses;
KW clone pMO-218.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..21 /label= Signal_peptide
FT Protein 22..466 /label= Human_mature_chitinase_protein
FT Domain 418..466 /label= Chitin_binding_domain
FT Region 1..373 /note= "Region with triacetylchitotriose hydrolysing activity"
XX
PN WO200123430-A2.
XX
PD 05-APR-2001.
XX
PF 28-SEP-2000; 2000WO-US26960.
XX
PR 30-SEP-1999; 99US-0409918.
XX
PA (ICOS-) ICOS CORP.
XX
PI Allison DS, Dietsch GN, Gray PW, Shaw KD, Steiner BH;
XX
DR WPI; 2001-266141/27.
DR N-PSDB; AAD03759.
XX
PT Novel chitinase immunoglobulin fusion product, useful for treating
PT fungal infections and reducing the amount of a non-chitinase antifungal
PT agent needed for the treatment -
XX
PS Claim 1; Page 32-33; 39pp; English.
XX
CC The present invention relates to a chitinase immunoglobulin (Ig) fusion
CC product, comprising a human chitinase fused to at least a portion of an
CC immunoglobulin chain. The fusion product is useful for treating fungal
CC infections (mycoses) such as candidiasis, aspergillosis, blastomycosis,
CC coccidioidomycosis, paracoccidioidomycosis, histoplasmosis, mucormycosis,
CC cryptococcosis, chromoblastomycosis, sporotrichosis and dermatophytoses.
CC The fusion protein is useful for reducing the amount of non-chitinase
CC antifungal agent needed to exert an antifungal activity. The fusion
CC protein is also useful for preparing
CC a medicament for the prophylactic or therapeutic treatment of fungal
CC infections. Chitinase immunoglobulin fusion product has unexpectedly
CC improved serum half-life and formulation properties.
CC The present sequence is human chitinase protein from clone pMO-218.
CC Chitinase enzyme degrades chitin which is a homopolymer of
CC beta-(1,4)-linked N-acetylglucosamine residues.
XX
SQ Sequence 466 AA;


```

RESULT 11
AAW40260
XX AAW40260 standard; Protein; 466 AA.
XX
XX AAW40260;
XX
DT 15-JUN-1998 (first entry)
XX
DE Human chitinase protein from clone MO-13B.
XX
KW Chitinase; human; fungal infection; immunogen; diagnosis; treatment;
KW Gaucher's disease; transgenic; detection; hybridisation; antifungal;
KW rheumatoid arthritis; overexpression; extracellular matrix.
XX
OS Homo sapiens.
XX
XX
FH Key Location/Qualifiers
FT Peptide 1..21
FT Protein /label= signal
FT /label= chitinase
FT /note= "from clone MO-13B"
XX
XX
PN MO9747752-A1.
XX
XX
XX 18-DEC-1997.
XX
XX 16-JUN-1997; 97WO-US10460.
XX
XX 14-JUN-1996; 96US-0663618.
XX
XX (ICOS-) ICOS CORP.
XX
XX Gray PW;
XX
XX WPI; 1998-052316/05.
XX N-PSDB; AAV10436.
XX
XX Nucleic acids encoding human chitinase - useful as antifungal
XX agents, especially in combination with other antifungals
XX
XX Claim 7; Page 44-45; 63pp; English.
XX
XX This sequence represents a novel human chitinase isolated from clone
XX MO-13B. Chitinases are useful for treating or preventing fungal infection
XX and as immunogens for generating antibodies which are used to purify,
XX detect and quantify chitinases, e.g. for diagnosis of Gaucher's disease.
XX The nucleic acid sequence of the chitinase is also useful as a probe to
XX identify and isolate genomic DNA encoding chitinases or similar proteins,
XX or cells expressing them or to generate transgenic ('knockout') rodents.
XX It can also be used in hybridisation assays and to detect genetic
XX alterations in the chitinase gene related to disease. Agents that inhibit
XX this protein may be useful in treatment of Gaucher's disease and
XX rheumatoid arthritis, where overexpression of the protein can damage
XX the extracellular matrix. Chitinase also improves the activity of other
XX antifungal agents and may allow a reduction in the dose of such agents,
XX and thus of their side effects.
XX
XX Sequence 466 AA;
XX
Query Match 52.6%; Score 1317; DB 19; Length 466;
Best Local Similarity 52.0%; Pred. No. 3.6e-111;
Matches 239; Conservative 70; Mismatches 127; Indels 24; Gaps 3;
QY 2 QLTCTYTNWAQYRPGLRFPDNDICLCTHLYIYAFAGQNNETIENWDTLYQAFNG 61
DB 23 KLVCTYTNWAQYRQGFARLPKDLDESCLCTHLYIYAFAGMTNQLSTWENDELYQEFNG 82
QY 62 LKNKNSQLKTLAIGGNWFTAPFTAMVSTPENRQFTFVSVKFLQYDFDGLDPEWYVP 121
DB 83 LKQNFPLKTLAIGGNWFTSTQFTDMVATANNRQIFVNSAIRFLKRYDFDGLDWEYVP 142
QY 122 GSRGSPQDKHLFTVLVQEMREAFQEAQINKPRLMVTAAVAAGISNIQSGYEIPQLSQ 181

```

```

143 GSQGSAPVDKERFTTLVQDLANAFQOEATSGKERLLLSAAVPAGQTYVDAGYVDKIAQ 202
182 YLDYIHWMTYDLHGSWEQYTGNSPLYKYPTDTGSNAVILNVVYVNVYKNDGAPAEKLIIV 241
203 NLDVFNLMAYDFHGSWEKVTGHSPLYKQESGAAASLNVDAAVQOQMLQKGTPAKLIIL 262
242 GPFTYGHNFILSNPNTGIGAPTSGAGPAGPYAKESGIWAYVEICTFLKNGATQGWDAPO 301
263 GNPTVGRSFTLASSSDTRVGAPATGSGTGPFTKEGMLAYVEVCSW--KGATKQRIQDO 320
302 EYPIAYQGVVWYGVYDNIKSFDIKAWLKHKNKFGGAMVWALDLDFTGTCNQGKPEPLIST 361
321 KVPYIFRDNQWVGFDDVESFKTKVSLKQKGLGGAMVWALDLDDEAGFSNCGRYPLIQT 380
362 LKKAL-----GLOSASCTAPAOPIETAPSGSGSGSGSGSGSGSGSGSGSGSGSGGFCFANVRAN 414
381 LROELSLPYLPSTGTPLEVPKGPQSEP-----EHGPPSPGQDTFCQKAD 425
415 GLYPVANNRNFHVCVNGVTYQNCQAGLVFDTSCDCCNW 454
426 GLYPNPRSSPYSACAGRLFOQSCPTGLVFSNSCKCCTW 465

RESULT 12
AAW42426
ID AAW42426 standard; Protein; 466 AA.
XX
XX AAW42426;
XX
DT 10-DEC-1999 (first entry)
XX
DE MO-13B clone of human Chitinase, amino acid sequence.
XX
KW chitin; fungal infection; immunocompromised; AIDS; chemotherapy;
KW organ transplant; parasite; chitin-binding; allele; vector;
KW truncated protein; chitin binding region.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..21
FT Protein /label= Signal_peptide
FT /label= Mature_protein
XX
XX WO9946390-A1.
XX
XX 16-SEP-1999.
XX
XX 12-MAR-1999; 99WO-US05343.
XX
XX 12-MAR-1998; 98US-0039198.
XX (ICOS-) ICOS CORP.
XX
XX Gray PW, Tjoelker LW;
XX
XX WPI; 1999-551417/46.
XX N-PSDB; AA221848.
XX
XX Novel chitin-binding fragments of human chitinase used to treat fungal
XX infections in animals
XX
XX Example 1; Page 62-64; 83pp; English.
XX
XX This is the amino acid sequence of an allelic form of the human
XX chitinase enzyme, which is capable of degrading Chitin (a linear
XX homo polymer of beta-1,4-linked N-acetylglucosamine residues).
XX The sequence also includes a Chitin binding peptide region, which lacks
XX any chitinase activity, but which can be used to target anti-fungal
XX agents to chitin containing fungal cells.
XX Chitinase fragments can be used to screen for proteins or other

```

CC molecules that specifically bind to the chitin-binding domain of human
CC chitinase or that modulate its activity. These compounds are useful for
CC immunization, as well as for purifying chitinase, as well as for
CC detection and quantification of chitinase. Polynucleotide fragments of
CC the invention are useful as a source of probes and primers, and to
CC express the proteins recombinantly. The chitinase fragments, when
CC conjugated to antifungal compounds, are used to treat animals,
CC especially humans, infected with chitin-containing parasites such as
CC fungi. Fungal infection treated include candidiasis, aspergillosis,
CC coccidioidomycosis, blastomycosis, paracoccidioidomycosis,
CC mucormycosis, histoplasmosis, cryptococcosis, chromoblastomycosis,
CC sporotrichosis, and dermatophytoses.
CC Chitin can be degraded by the enzyme chitinase. Use of whole chitinase
CC protein for treating infections, especially fungal infections, is
CC problematic. In view of the increasing incidents of life-threatening
CC fungal infection in e.g. immunocompromised individuals, there exists a
CC need for identifying new compounds for treating fungal infection. The
CC chitin-binding fragments of the present invention provide this need.
XX
SQ Sequence 466 AA;

Query Match 52.6%; Score 1317; DB 20; Length 466;
Best Local Similarity 52.0%; Pred. No. 3.6e-111;
Matches 239; Conservative 70; Mismatches 127; Indels 24; Gaps 3;

QY 2 QLTCTFTNWAQYRPGLGFRMPDNIDPCLCTHLIYAFAGRNNEITTEWNDVTLQAFNG 61
DB 23 KLVCTFTNWAQYRQGEARLPKDLPSLCTHLIYAFAGMTNQLSTTEWDELTYQEFNG 82
QY 62 LKKNKSQLKTLTLLAIGWNFGTAPFTAMVSTPENROTFITSVIKRQYEFGLDFDWEYP 121
DB 83 LKKNPKLTLTLLAIGWNFGTAPFTAMVSTPENROTFITSVIKRQYEFGLDFDWEYP 142
QY 122 GSRGSPDQKHFTVLVQEMREAFQEAQINPKRLMVTAAVAAGISNIQSYEIPQLSQ 181
DB 83 LKKNPKLTLTLLAIGWNFGTAPFTAMVSTPENROTFITSVIKRQYEFGLDFDWEYP 142
QY 182 YLDYTHVMTYDLHSGWEGTGENSPLYKYPTDTGSMAYLVNDVYVMYWKDNGAPAEKLTIV 241
DB 203 NLDVFNLMAYDPHGSWEKVTGHNPSPLYKQESGAASLVNDAVQOVLQKTPASKLIL 262
QY 242 GPPTGHNFLNPSNTGIGTSGAGPAGPVAKSGIWAYEYICFTFLKNGATQGWDAPO 301
DB 263 GMPTYGRSFTLASSSDTRVAGPATSQGTGPGFTKGGMLAYEVCWS--KGATKQRIQDQ 320
QY 302 EYPYAVQGNVWGYDNIKSFDIAKWLKHNKFGGMWALDLDGTFCTNKGKPLIST 361
DB 321 KVPYIFRDQWGVDFVESFKTKVSLYKQKGLGGAMWALDLDGTFCTNKGKPLIST 380
QY 362 LKKAL-----GLOSASCTAPAQPIETITAPSGSGNGSGSGSGSGSGSGGFCVAVRAN 414
DB 381 LKQELSLPVLPSGTPELVKPGQSEP-----EHGPFSGQDTFCQKAD 425
QY 415 GLYPVANNRPAFWCHVNGVYQONQOAGLVFTSDCCNW 454
DB 426 GLYPNPRSSPYSCAAGRLFQSCPTGLVFNSSCKCTW 465

RESULT 13
AAE00433
XX
AC AAE00433;
XX
DT 19-JUN-2001 (first entry)
XX Human chitinase protein from clone pMO-13B.
XX Human; antifungal; chitinase; immunoglobulin; Ig; therapy;
KW fungal infection; candidiasis; aspergillosis; coccidioidomycosis;
KW blastomycosis; paracoccidioidomycosis; histoplasmosis; cryptococcosis;
KW chromoblastomycosis; sporotrichosis; mucormycosis; dermatophytoses;
KW clone pMO-13B.

XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH Peptide 1..21 /label= Signal_peptide
FT Protein 22..466 /label= Human_mature_chitinase_protein
FT Domain 418..466 /label= Chitin_binding_domain
FT Region 1..373 /note= "Region with trisacetylchitotriose
FT hydrolysing activity"
XX
XX WO200123430-A2.
XX
XX 05-APR-2001.
XX
XX 28-SEP-2000; 2000WO-US26960.
XX
XX 30-SEP-1999; 99US-0409918.
XX
XX (ICOS-) ICOS CORP.
XX
XX Allison DS, Dietsch GN, Gray PW, Shaw KD, Steiner BH;
XX
XX WPI: 2001-266141/27.
XX N-PSDB; AAD03760.
XX
XX Novel chitinase immunoglobulin fusion product, useful for treating
XX fungal infections and reducing the amount of a non-chitinase antifungal
XX agent needed for the treatment -
XX
XX Claim 1; Page 36-38; 39pp; English.
XX
XX The present invention relates to a chitinase immunoglobulin (Ig) fusion
XX product, comprising a human chitinase fused to at least a portion of an
XX immunoglobulin chain. The fusion product is useful for treating fungal
XX infections (mycoses) such as candidiasis, aspergillosis, blastomycosis,
XX coccidioidomycosis, paracoccidioidomycosis, histoplasmosis, mucormycosis,
XX cryptococcosis, chromoblastomycosis, sporotrichosis and dermatophytoses.
XX The fusion protein is useful for reducing the amount of non-chitinase
XX antifungal agent needed to exert an antifungal activity. The fusion
XX protein is also useful for preparing
XX a medicament for the prophylactic or therapeutic treatment of fungal
XX infections. Chitinase immunoglobulin fusion product has unexpectedly
XX improved serum half-life and formulation properties.
XX The present sequence is human chitinase protein from clone pMO-13B.
XX Chitinase enzyme degrades chitin which is a homopolymer of
XX beta-(1,4)-linked N-acetylglucosamine residues.
XX
XX Sequence 466 AA;

Query Match 52.6%; Score 1317; DB 22; Length 466;
Best Local Similarity 52.0%; Pred. No. 3.6e-111;
Matches 239; Conservative 70; Mismatches 127; Indels 24; Gaps 3;

QY 2 QLTCTFTNWAQYRPGLGFRMPDNIDPCLCTHLIYAFAGRNNEITTEWNDVTLQAFNG 61
DB 23 KLVCTFTNWAQYRQGEARLPKDLPSLCTHLIYAFAGMTNQLSTTEWDELTYQEFNG 82
QY 62 LKKNKSQLKTLTLLAIGWNFGTAPFTAMVSTPENROTFITSVIKRQYEFGLDFDWEYP 121
DB 83 LKKNPKLTLTLLAIGWNFGTAPFTAMVSTPENROTFITSVIKRQYEFGLDFDWEYP 142
QY 122 GSRGSPDQKHFTVLVQEMREAFQEAQINPKRLMVTAAVAAGISNIQSYEIPQLSQ 181
DB 143 GSGSPAVDKERFTTLVQDLANAFQEAQTSKGERLLLSAAVPAQTYVDAGYEVDKIAQ 202
QY 182 YLDYTHVMTYDLHSGWEGTGENSPLYKYPTDTGSMAYLVNDVYVMYWKDNGAPAEKLTIV 241
DB 203 NLDVFNLMAYDPHGSWEKVTGHNPSPLYKQESGAASLVNDAVQOVLQKTPASKLIL 262

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2003, 21:01:56 ; Search time 17.5 Seconds
(without alignments)
2850.818 Million cell updates/sec

Title: US-10-004-219B-14

Perfect score: 2506

Sequence: 1 YOLTCYFTNWAQYRGLGRF.....QQNCQAGLVFTSDCCNWA 455

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 424699 seqs, 109646833 residues

Total number of hits satisfying chosen parameters: 424699

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US08_NEW PUB.pbp.*
- 2: /cgn2_6/ptodata/1/pubpaa/ECT NEW PUB.pbp.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW PUB.pbp.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pbp.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW PUB.pbp.*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pbp.*
- 7: /cgn2_6/ptodata/1/pubpaa/ECTUS_PUBCOMB.pbp.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pbp.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_NEW PUB.pbp.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pbp.*
- 11: /cgn2_6/ptodata/1/pubpaa/US10_NEW PUB.pbp.*
- 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pbp.*
- 13: /cgn2_6/ptodata/1/pubpaa/US60_NEW PUB.pbp.*
- 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pbp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2506	100.0	455	9	US-10-004-219B-14
2	2506	100.0	476	9	US-10-004-219B-1
3	2089.5	83.4	452	9	US-10-004-219B-9
4	2089.5	83.4	473	9	US-10-004-219B-4
5	1323	52.8	445	9	US-10-004-219B-10
6	976.5	39.0	383	10	US-10-097-340-45
7	968.5	38.6	383	10	US-09-459-749D-17
8	711.5	28.4	490	9	US-10-218-743-41
9	711.5	28.4	509	9	US-10-218-743-35
10	711.5	28.4	509	9	US-10-218-743-38
11	708.5	28.3	536	9	US-10-218-743-21
12	708.5	28.3	555	9	US-10-218-743-15
13	708.5	28.3	555	9	US-10-218-743-18
14	472	18.8	371	9	US-09-923-844B-2
15	285.5	11.4	376	10	US-09-748-033-3
16	196	7.4	170	9	US-10-218-743-44
17	104.5	4.2	711	9	US-10-071-485-90
18	104	4.2	309	9	US-10-304-928-10
19	98.5	3.9	1385	10	US-09-738-363-2

20	98	3.9	750	9	US-10-046-433-2	Sequence 2, Appli
21	98	3.9	1001	9	US-10-046-433-40	Sequence 40, Appl
22	98	3.9	1013	9	US-10-028-072-38	Sequence 38, Appl
23	98	3.9	1013	9	US-10-121-049-38	Sequence 38, Appl
24	98	3.9	1013	9	US-10-123-904-38	Sequence 38, Appl
25	98	3.9	1013	9	US-10-140-470-38	Sequence 38, Appl
26	98	3.9	1013	9	US-10-175-746-38	Sequence 38, Appl
27	98	3.9	1013	9	US-10-176-918-38	Sequence 38, Appl
28	98	3.9	1013	9	US-10-176-921-38	Sequence 38, Appl
29	98	3.9	1013	9	US-10-137-865-38	Sequence 38, Appl
30	98	3.9	1013	9	US-10-140-474-38	Sequence 38, Appl
31	98	3.9	1013	9	US-10-142-431-38	Sequence 38, Appl
32	98	3.9	1013	9	US-10-143-114-38	Sequence 38, Appl
33	98	3.9	1013	9	US-10-140-002-38	Sequence 38, Appl
34	98	3.9	1013	9	US-10-142-419-38	Sequence 38, Appl
35	98	3.9	1013	9	US-10-123-262-38	Sequence 38, Appl
36	98	3.9	1013	9	US-10-142-423-38	Sequence 38, Appl
37	98	3.9	1013	9	US-10-121-050-38	Sequence 38, Appl
38	98	3.9	1013	9	US-10-141-755-38	Sequence 38, Appl
39	98	3.9	1013	9	US-10-143-032-38	Sequence 38, Appl
40	98	3.9	1013	9	US-10-123-108-38	Sequence 38, Appl
41	98	3.9	1013	9	US-10-123-236-38	Sequence 38, Appl
42	98	3.9	1013	9	US-10-123-261-38	Sequence 38, Appl
43	98	3.9	1013	9	US-10-140-921-38	Sequence 38, Appl
44	98	3.9	1013	9	US-10-140-928-38	Sequence 38, Appl
45	98	3.9	1013	9	US-10-121-045-38	Sequence 38, Appl

ALIGNMENTS

RESULT 1
US-10-004-219B-14
; Sequence 14, Application US/10004219B
; Publication No. US20030087414A1
; GENERAL INFORMATION:
; APPLICANT: Macrozyme
; APPLICANT: Aerts, Johannes M.F.G.
; APPLICANT: Boot, Rolf G.
; TITLE OF INVENTION: A mammalian mucinase, its recombinant production, and its use in therapy or prophylaxis against diseases in which mucus is involved or infection diseases
; TITLE OF INVENTION: its use in therapy or prophylaxis against diseases in which mucus is involved or infection diseases
; FILE REFERENCE: 2183-5136US
; CURRENT APPLICATION NUMBER: US/10/004,219B
; CURRENT FILING DATE: 2001-11-02
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1)..(455)
; OTHER INFORMATION: /note="Human AMCase"
US-10-004-219B-14

Query Match 100.0%; Score 2506; DB 9; Length 455;
Best Local Similarity 100.0%; Pred. No. 6e-204; Mismatches 0; Indels 0; Gaps 0;
Matches 455; Conservative 0;

Qy	1	YOLTCYFTNWAQYRGLGRFMPDNIDPCLCTHLIYAFAGRONNEITTIENWNTVLYQAFN 60
Db	1	YOLTCYFTNWAQYRGLGRFMPDNIDPCLCTHLIYAFAGRONNEITTIENWNTVLYQAFN 60
Qy	61	GLKKNKSQLKTLAIGGWNFGTAPMTAMVSTPENRQTITSVIKFLRQYFDFGLDFDWEY 120
Db	61	GLKKNKSQLKTLAIGGWNFGTAPMTAMVSTPENRQTITSVIKFLRQYFDFGLDFDWEY 120
Qy	121	PGSRGSPQDKHLFTVLVOEMREAEQAKQINKPRLMVTAAVAAGISNIQSGYEIPQLS 180
Db	121	PGSRGSPQDKHLFTVLVOEMREAEQAKQINKPRLMVTAAVAAGISNIQSGYEIPQLS 180

RESULT 4

US-10-004-219B-4
; Sequence 4, Application US/10004219B
; Publication No. US2003008741A1
; GENERAL INFORMATION:
; APPLICANT: Macrozyme
; APPLICANT: Aerts, Johannes M.F.G.
; APPLICANT: Boot, Rolf G.
; TITLE OF INVENTION: A mammalian mucinase, its recombinant production, and
; TITLE OF INVENTION: its use in therapy or prophylaxis against diseases in
; TITLE OF INVENTION: which mucus is involved or infection diseases
; FILE REFERENCE: 2183-5136US
; CURRENT APPLICATION NUMBER: US/10/004,219B
; CURRENT FILING DATE: 2001-11-02
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 473
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: mouse AWCae
; OTHER INFORMATION: amino acid sequence deduced from cDNA sequence
US-10-004-219B-4

Query Match 83.4%; Score 2089.5; DB 9; Length 473;
Best Local Similarity 81.3%; Pred. No. 1.2e-168;
Matches 369; Conservative 40; Mismatches 42; Indels 3; Gaps 1;
QY 1 YQLTCYFTNWAQYRPGLGFMFDNIDPCLCTHLIYAFAGQNNETIIEWNDVTLYQAFN 60
DB 22 YNLICFTNWAQYRPGLGSKFDNIDPCLCTHLIYAFAGQNNETIIEWNDVTLYKAFN 81
QY 61 GLKKNLSQKLTLLAIGGWNFGTAPFTAMYSTENROTEITSVKFLROYEFDGLDFDWEY 120
DB 82 DLKRNRSKLTLLAIGGWNFGTAPFTAMYSTENROTEITSVKFLROYEFDGLDFDWEY 141
QY 121 PGSRSGPPQDKHLFTVLVQEMREAFQEAQKQINPKRLMTAAVAGISNIQSGYEIPQLS 180
DB 142 PGSRSGPPQDKHLFTVLVQEMREAFQEAQKQINPKRLMTAAVAGISNIQSGYEIPQLS 201
QY 181 QYLDYIHVMTYDLHGSWEGYTGNSPLYKYPTDTGNSNAYLVNDVYVWYKNGAPAEKLI 240
DB 202 KYLDFTHVMTYDLHGSWEGYTGNSPLYKYPTDTGNSNAYLVNDVYVWYKNGAPAEKLI 261
QY 241 VGPFTYGHNFILSNPNTGIGAPTSAGAPYAKESGWAYYEICTFLKNGATQGWADAP 300
DB 262 VGPFTYGHNFILSNPNTGIGAPTSAGAPYAKESGWAYYEICTFLKNGATQGWADAP 321
QY 301 QYVPYAYQGVVWYGVYDNIKSFDIKQWLKHNKFGGAMVWALDLDFTGTCNQGKPEPLIS 360
DB 322 QYVPYAYKANEMLYGVYDNIKSFSVKAQWLKQNNFGGAMVWALDLDFTGTCNQGKPEPLIS 381
QY 361 TLKALGLQASCTAPAPQIEPTAPSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSG 420
DB 382 TLKALGISTEGCTAPDVPSEPTTTPP---GSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSG 438
QY 421 NNENAFWHCVNGVYQONCOAGLVFDTSCDCCNW 454
DB 439 DORNAPWQINGITYQOHCOAGLVFDTSCNCCNW 472

RESULT 5

US-10-004-219B-10
; Sequence 10, Application US/10004219B
; Publication No. US2003008741A1
; GENERAL INFORMATION:
; APPLICANT: Macrozyme
; APPLICANT: Aerts, Johannes M.F.G.
; APPLICANT: Boot, Rolf G.
; TITLE OF INVENTION: A mammalian mucinase, its recombinant production, and
; TITLE OF INVENTION: its use in therapy or prophylaxis against diseases in
; TITLE OF INVENTION: which mucus is involved or infection diseases

FILE REFERENCE: 2183-5136US
; CURRENT APPLICATION NUMBER: US/10/004,219B
; CURRENT FILING DATE: 2001-11-02
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1)...(445)
; OTHER INFORMATION: /notes="Human chitotriosidase"
US-10-004-219B-10

Query Match 52.8%; Score 1323; DB 9; Length 445;
Best Local Similarity 52.2%; Pred. No. 8.5e-104;
Matches 240; Conservative 70; Mismatches 126; Indels 24; Gaps 3;
QY 2 QLTCTYFTNWAQYRPGLGFMFDNIDPCLCTHLIYAFAGQNNETIIEWNDVTLYQAFNG 61
DB 2 KLVCTYFTNWAQYRQGEARFLPKDLPSLCTHLIYAFAGWTHQLSTTEWNDVTLYQEFNG 61
QY 62 LKNKNSQLKTLAIGGWNFGTAPFTAMYSTENROTEITSVKFLROYEFDGLDFDWEY 121
DB 62 LKNKNSQLKTLAIGGWNFGTAPFTAMYSTENROTEITSVKFLROYEFDGLDFDWEY 121
QY 122 GSRSGPPQDKHLFTVLVQEMREAFQEAQKQINPKRLMTAAVAGISNIQSGYEIPQLSQ 181
DB 122 GSGSGPVDKFRFTVLVQDLANAFQEAQTSCKERLLLSAAVPAGQTYVDAGYVDKIAQ 181
QY 182 YLDYIHVMTYDLHGSWEGYTGNSPLYKYPTDTGNSNAYLVNDVYVWYKNGAPAEKLI 241
DB 182 NLDVNLMAVDYFHGSWEKVTGHSNPLYKQESGAAASLNVDAVQWLQKTPASKLIL 241
QY 242 GPTTYGHNFILSNPNTGIGAPTSAGAPYAKESGWAYYEICTFLKNGATQGWADAP 301
DB 242 GMPTYGRSFTLASSSDTRVGAPATGSGTPGPTKEGGLAYYEVCWS--KGATKQRIQDQ 299
QY 302 EYVPYAYQGVVWYGVYDNIKSFDIKQWLKHNKFGGAMVWALDLDFTGTCNQGKPEPLIS 361
DB 300 KVPYIFRDNQWVGFDVDFSTKYSYLKQKGLGGAMVWALDLDFTGTCNQGKPEPLIS 359
QY 362 LKAL-----GLQASCTAPAPQIEPTAPSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSG 414
DB 360 LQELSLPLPSGTPELEVPKPGQSEP-----EHGSPSGQDTFCQKAD 404
QY 415 GLYPVANNRPAFWHCVNGVYQONCOAGLVFDTSCDCCNW 454
DB 405 GLYPNPRSSFSYSCAAGRLFOQSCPTGLVFSNCKCCTW 444

RESULT 6

US-10-097-340-45
; Sequence 45, Application US/10097340
; Publication No. US20030087250A1
; GENERAL INFORMATION:
; APPLICANT: John MONAHAN
; APPLICANT: Marijula GANNARAPU
; APPLICANT: Sebastian HOERSCH
; APPLICANT: Shubhangi KAMATKAR
; APPLICANT: Steve G. KOVATS
; APPLICANT: Rachel E. MEYERS
; APPLICANT: Michael MORRISSEY
; APPLICANT: Peter OLANDT
; APPLICANT: Ami SEN
; APPLICANT: Peter VEIBY
; APPLICANT: Gordon B. MILLS
; APPLICANT: Robert C. BAST, Jr.
; APPLICANT: Karen LU
; APPLICANT: Rosemarie SCHMANDT
; APPLICANT: Xumei ZHAO
; APPLICANT: Karen GLATT

; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
 ; FILE REFERENCE: MRI-030
 ; CURRENT APPLICATION NUMBER: US/10/097,340
 ; CURRENT FILING DATE: 2002-03-14
 ; PRIOR APPLICATION NUMBER: 60/276,025
 ; PRIOR FILING DATE: 2001-03-14
 ; PRIOR APPLICATION NUMBER: 60/325,149
 ; PRIOR FILING DATE: 2001-09-26
 ; PRIOR APPLICATION NUMBER: 60/276,026
 ; PRIOR FILING DATE: 2001-03-14
 ; PRIOR APPLICATION NUMBER: 60/324,967
 ; PRIOR FILING DATE: 2001/09/26
 ; PRIOR APPLICATION NUMBER: 60/311,732
 ; PRIOR FILING DATE: 2001-08-10
 ; PRIOR APPLICATION NUMBER: 60/325,102
 ; PRIOR FILING DATE: 2001-09-26
 ; PRIOR APPLICATION NUMBER: 60/323,580
 ; PRIOR FILING DATE: 2001-09-19
 ; NUMBER OF SEQ ID NOS: 363
 ; SOFTWARE: Fast-Seq for Windows Version 4.0
 ; SEQ ID NO 45
 ; * TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-097-340-45

Query Match 39.0%; Score 976.5; DB 9; Length 383;
 Best Local Similarity 51.2%; Pred. No. 1.5e-74;
 Matches 188; Conservative 62; Mismatches 108; Indels 9; Gaps 5;
 QY 1 YQLTCYFTNWAQYRGLRPMNDPCLCTHLYAFAGRQNNETITIEWNDVTLQAFN 60
 Db 22 YKLVCTYTSWSQYREGDSCFPDADPFLCTHLYIFSFANISNNEIDTLEWNDVTLQAFN 81
 QY 61 GLKNKNSQLKTLAIGGNWFGTAPMTAVSTPENQRTFTSVIKFLQYEFDFGLDPWEY 120
 Db 82 TLKRNPNLKTLLSVGNWFGSQRFSKIASNTQSRRTFKSVPPFLRTHGFDGLDLAWLY 141
 QY 121 PGSRGSPQDKHLFTVLVQEMREAFQEAQKINKPRLMVTAAVAAGISNIQSYEIPQLS 180
 Db 142 PGR-----DKRHLLTLVKEMKAEFVREALP-GTERLLSGAVSAGKVALDRGYDIAQIS 195
 QY 181 QYLDYIHVMYDHLGWSWEGYTGNSPLKYKPTDTGNSNAYLVNVDYVMYWKDNGAPAEKLI 240
 Db 196 QHLDFTISIMTYDFHGAWRGTTGHSPLFRQGEDASDRFSNTDYAVGYMLRGLGAPASKLV 255
 QY 241 VGPTTYGHNFILSPNNTGIGAPTSGAGPAGPYAKESGIWAYVEICTFLKNGATQGWDPAP 300
 Db 256 MGIPTFGRSFTLAS-SETGVGAPISGPGIPGRFTKEAGTILAYVEICDFLR-GATVHRTLG 313
 QY 301 QEVYAYQGNVWGYDNIKSFDPKQWLKHNKFGGMVWAIIDDDTGTGFCNQG-KFPLI 359
 Db 314 QQVYATKGNQWGYDDQESVKSQVYLKDRQLAGAMVWALDLDLDFQSGFCGQDLFFPLT 373
 QY 360 STLKAL 366
 Db 374 NAIKDAL 380

RESULT 7
 US-09-459-749D-17
 ; Sequence 17, Application US/09459749D
 ; Patent No. US20020136716A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Millis, Albert J. T.
 ; TITLE OF INVENTION: Compositions and Methods For Altering Cell Migration
 ; FILE REFERENCE: 0794.016A
 ; CURRENT APPLICATION NUMBER: US/09/459,749D
 ; CURRENT FILING DATE: 1999-12-10
 ; PRIOR APPLICATION NUMBER: 60/111,856
 ; PRIOR FILING DATE: 1998-12-11
 ; NUMBER OF SEQ ID NOS: 17

; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 17
 ; LENGTH: 383
 ; TYPE: PRT
 ; ORGANISM: Sus scrofa
 US-09-459-749D-17
 Query Match 38.6%; Score 968.5; DB 10; Length 383;
 Best Local Similarity 51.5%; Pred. No. 7.2e-74;
 Matches 189; Conservative 57; Mismatches 112; Indels 9; Gaps 5;
 QY 1 YQLTCYFTNWAQYRGLRPMNDPCLCTHLYAFAGRQNNETITIEWNDVTLQAFN 60
 Db 22 YKLVCTYTSWSQYREGDSCFPDADPFLCTHLYIFSFANISNNEIDTLEWNDVTLQAFN 81
 QY 61 GLKNKNSQLKTLAIGGNWFGTAPMTAVSTPENQRTFTSVIKFLQYEFDFGLDPWEY 120
 Db 82 TLKRNPNLKTLLSVGNWFGSQRFSKIASNTQSRRTFKSVPPFLRTHGFDGLDLAWIS 141
 QY 121 PGSRGSPQDKHLFTVLVQEMREAFQEAQKINKPRLMVTAAVAAGISNIQSYEIPQLS 180
 Db 142 PGR-----DKRHLLTLVKEMKAEFVREALP-GTERLLSGAVSAGKVALDRGYDIAQIS 195
 QY 181 QYLDYIHVMYDHLGWSWEGYTGNSPLKYKPTDTGNSNAYLVNVDYVMYWKDNGAPAEKLI 240
 Db 196 QHLDFTISIMTYDFHGAWRGTTGHSPLFRQGEDASDRFSNADYAVSVYLRGLGAPASKLV 255
 QY 241 VGPTTYGHNFILSPNNTGIGAPTSGAGPAGPYAKESGIWAYVEICTFLKNGATQGWDPAP 300
 Db 256 MGIPTFGRSFTLAS-SKTDVGAPASGPGIPGRFTKEGILAYVEICDFLR-GATVRRPLG 313
 QY 301 QEVYAYQGNVWGYDNIKSPDKQWLKHNKFGGMVWAIIDDDTGTGFCNQG-KFPLI 359
 Db 314 QQVYATKGNQWGYDDQESVKNKAKILKSRQLAGAMVWTLDDDFRGNFCQGNLFFPLT 373
 QY 360 STLKAL 366
 Db 374 NAIKDV 380

RESULT 8
 US-10-218-743-41
 ; Sequence 41, Application US/10218743
 ; Publication No. US20030096779A1
 ; GENERAL INFORMATION:
 ; APPLICANT: McCall, Catherine A.
 ; APPLICANT: Hunter, Shirley Wu
 ; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
 ; TITLE OF INVENTION: AND USES THEREOF
 ; FILE REFERENCE: AL-2-C3
 ; CURRENT APPLICATION NUMBER: US/10/218,743
 ; CURRENT FILING DATE: 2002-08-13
 ; PRIOR APPLICATION NUMBER: US/09/292,225
 ; PRIOR FILING DATE: 1999-04-15
 ; PRIOR APPLICATION NUMBER: 60/098,909
 ; PRIOR FILING DATE: 1998-09-02
 ; PRIOR APPLICATION NUMBER: 60/085,295
 ; PRIOR FILING DATE: 1998-05-13
 ; PRIOR APPLICATION NUMBER: 60/098,565
 ; PRIOR FILING DATE: 1998-04-17
 ; PRIOR APPLICATION NUMBER: 09/062,013
 ; PRIOR FILING DATE: 1998-04-17
 ; NUMBER OF SEQ ID NOS: 49
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 41
 ; LENGTH: 490
 ; TYPE: PRT
 ; ORGANISM: Dermatophagoides farinae
 US-10-218-743-41
 Query Match 28.4%; Score 711.5; DB 9; Length 490;
 Best Local Similarity 36.4%; Pred. No. 5.8e-52;

Matches 148; Conservative 74; Mismatches 152; Indels 33; Gaps 10;

QY 2 QLTCTYTNWAQYRPGLGFRMPDNIDPCLCTHLIYAFAGRQNNETIENW----- 51
Db 15 RIVCYVGTWSVYHK-VDPYTIEDIDPFKCTHLMYGFAKIDYKYTIQVDFPDDNNHNSW 73

QY 52 DVTLYQAFNGLKNKNSQLKTLAIGWNFGTAPFTAMVSTPENROTFTITSVKFLRQYEF 111
Db 74 EKHGVERFNNLRKKNPELTMTISLGGWYEGSEKYSDMAANPTYRQFVQSVDLDFLOEYKF 133

QY 112 DGLDFDWEYPGSR-GSPPODKHLFTVLVQEMREAFQEAQINPKRLMVTAAVAAGISNI 170
Db 134 DGLDLDWEYPGSRGLGNPKIDKQNYLTTLVRELKEAFEPFG-----YLLTAASVPGKDKI 186

QY 171 QSGYEIPOLSOYLDYIHVMTYDLHGSWEGYTGENSPLYKYPTDGTGS-NAYLNVDVYMNW 229
Db 187 DVAYELKELNQLFDWMNVMTYDHGWNVFVGHNAPLYKRPDETDELHTYFNVTMTHY 246

QY 230 KONGAPAEKLIYGVFTYGHNFILSNPSTGIGAPTSAGPAGPYAKESGIWAYEIC-TF 288
Db 247 LNNGATRDKLNVGPPYGYGRAWSDRSKVKLGDPKAGMSPPGFITGEEGVLSYIELCQLF 306

QY 289 LKNGATQGWADAPQEVYAYQGNVWGYDNKISFQDKAQLKHNKFGGAMVWALDDDFTG 348
Db 307 QKEEWHIOYDEYNAPYGYNDKIWGYDDLASISCKLAFKELGVLGSGVMWISLENDDFKG 366

QY 349 TFCNKGKPELSTLKKALG-----LOSASCTAPAOPIEPITAA 386
Db 367 -HCGP-KYPLLKNVHNMINGDEKNSYECLLGPSTTTPT-PTTPSTTS 410

RESULT 9

US-10-218-743-35
; Sequence 35, Application US/10218743
; Publication No. US20030096779A1
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
; FILE REFERENCE: AL-2-C3
; CURRENT FILING DATE: 2002-08-13
; PRIOR FILING DATE: 1998-09-02
; PRIOR FILING DATE: 1998-05-13
; PRIOR FILING DATE: 1998-04-15
; PRIOR FILING DATE: 1998-04-17
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 35
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Dermatophagoides farinae
US-10-218-743-35

Query Match 28.4%; Score 711.5; DB 9; Length 509;
Best Local Similarity 36.4%; Pred. No. 6.1e-52;
Matches 148; Conservative 74; Mismatches 152; Indels 33; Gaps 10;

QY 2 QLTCTYTNWAQYRPGLGFRMPDNIDPCLCTHLIYAFAGRQNNETIENW----- 51
Db 34 RIVCYVGTWSVYHK-VDPYTIEDIDPFKCTHLMYGFAKIDYKYTIQVDFPDDNNHNSW 92

QY 52 DVTLYQAFNGLKNKNSQLKTLAIGWNFGTAPFTAMVSTPENROTFTITSVKFLRQYEF 111
Db 93 EKHGVERFNNLRKKNPELTMTISLGGWYEGSEKYSDMAANPTYRQFVQSVDLDFLOEYKF 152

QY 112 DGLDFDWEYPGSR-GSPPODKHLFTVLVQEMREAFQEAQINPKRLMVTAAVAAGISNI 170
Db 153 DGLDLDWEYPGSRGLGNPKIDKQNYLTTLVRELKEAFEPFG-----YLLTAASVPGKDKI 205

QY 171 QSGYEIPOLSOYLDYIHVMTYDLHGSWEGYTGENSPLYKYPTDGTGS-NAYLNVDVYMNW 229
Db 206 DVAYELKELNQLFDWMNVMTYDHGWNVFVGHNAPLYKRPDETDELHTYFNVTMTHY 265

QY 230 KONGAPAEKLIYGVFTYGHNFILSNPSTGIGAPTSAGPAGPYAKESGIWAYEIC-TF 288
Db 266 LNNGATRDKLNVGPPYGYGRAWSDRSKVKLGDPKAGMSPPGFITGEEGVLSYIELCQLF 325

QY 289 LKNGATQGWADAPQEVYAYQGNVWGYDNKISFQDKAQLKHNKFGGAMVWALDDDFTG 348
Db 326 QKEEWHIOYDEYNAPYGYNDKIWGYDDLASISCKLAFKELGVLGSGVMWISLENDDFKG 385

QY 349 TFCNKGKPELSTLKKALG-----LOSASCTAPAOPIEPITAA 386
Db 367 -HCGP-KYPLLKNVHNMINGDEKNSYECLLGPSTTTPT-PTTPSTTS 429

RESULT 10

US-10-218-743-38
; Sequence 38, Application US/10218743
; Publication No. US20030096779A1
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
; FILE REFERENCE: AL-2-C3
; CURRENT FILING DATE: 2002-08-13
; PRIOR FILING DATE: 1998-09-02
; PRIOR FILING DATE: 1999-04-15
; PRIOR FILING DATE: 1998-04-15
; PRIOR FILING DATE: 1998-09-02
; PRIOR FILING DATE: 1998-05-13
; PRIOR FILING DATE: 1998-04-17
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 38
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Dermatophagoides farinae
US-10-218-743-38

Query Match 28.4%; Score 711.5; DB 9; Length 509;
Best Local Similarity 36.4%; Pred. No. 6.1e-52;
Matches 148; Conservative 74; Mismatches 152; Indels 33; Gaps 10;

QY 2 QLTCTYTNWAQYRPGLGFRMPDNIDPCLCTHLIYAFAGRQNNETIENW----- 51
Db 34 RIVCYVGTWSVYHK-VDPYTIEDIDPFKCTHLMYGFAKIDYKYTIQVDFPDDNNHNSW 92

QY 52 DVTLYQAFNGLKNKNSQLKTLAIGWNFGTAPFTAMVSTPENROTFTITSVKFLRQYEF 111
Db 93 EKHGVERFNNLRKKNPELTMTISLGGWYEGSEKYSDMAANPTYRQFVQSVDLDFLOEYKF 152

QY 112 DGLDFDWEYPGSR-GSPPODKHLFTVLVQEMREAFQEAQINPKRLMVTAAVAAGISNI 170
Db 153 DGLDLDWEYPGSRGLGNPKIDKQNYLTTLVRELKEAFEPFG-----YLLTAASVPGKDKI 205

QY 171 QSGYEIPOLSOYLDYIHVMTYDLHGSWEGYTGENSPLYKYPTDGTGS-NAYLNVDVYMNW 229
Db 206 DVAYELKELNQLFDWMNVMTYDHGWNVFVGHNAPLYKRPDETDELHTYFNVTMTHY 265

APPLICANT: McCall, Catherine A.
APPLICANT: Hunter, Shirley Wu
APPLICANT: Weber, Eric R.
TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
FILE REFERENCE: AL-2-C3
CURRENT APPLICATION NUMBER: US/10/218,743
PRIORITY FILING DATE: 2002-08-13
PRIORITY APPLICATION NUMBER: US/09/292,225
PRIORITY FILING DATE: 1999-04-15
PRIORITY APPLICATION NUMBER: 60/098,909
PRIORITY FILING DATE: 1998-09-02
PRIORITY APPLICATION NUMBER: 60/085,295
PRIORITY FILING DATE: 1998-05-13
PRIORITY APPLICATION NUMBER: 60/098,565
PRIORITY FILING DATE: 1998-04-17
PRIORITY APPLICATION NUMBER: 09/062,013
PRIORITY FILING DATE: 1998-04-17
NUMBER OF SEQ ID NOS: 49
SOFTWARE: Patent in Ver. 2.0
SEQ ID NO 18
LENGTH: 555
TYPE: PRT
ORGANISM: Dermatophagoides farinae
US-10-218-743-18

Query Match 28.3%; Score 708.5; DB 9; Length 555;
Best Local Similarity 35.7%; Pred. No. 1.2e-51;
Matches 148; Conservative 77; Mismatches 154; Indels 35; Gaps 9;

QY 2 QLTCTFTWAQYRPGGRMPDNIDPCLCTHLIYAFAGRONNEITIEWN-----51
DB 34 RIVCYVGTWSVYHK-VDPYTIEDIDPFCFTHLMYGFADYKTYTQVDPYQDDNNHNSW 92

QY 52 DVTLYQAFNLGNKNSQLTLLAIGGNFCTAPFTAMVSTPENROFTITSVIFLQRYEP 111
DB 93 EKRGRYFNRLKLNKPELTMTLSLGGWYSGSEKISDMAANPTVROQFIQSVLDFLOEYKF 152

QY 112 DGLDFWMEYFGRS-GSPPODKHLFTVLVQEMREAFQEAQKINPKRLMTAAVAAGISNI 170
DB 153 DGLDLDFWMEYFGRSGLGNPKIDKQNYLALVRELKDAFEPHG-----YLLTAASPGKDKI 205

QY 171 QSGYEIPQLSOYLDYIHVMYDYLHGSWEGYTGNSPLKYPTDTGS-NAYLVNDVYMNW 229
DB 206 DRAYDIELKELNKLFDWNVMTYDYGWENFYGHNAPLYRPOEDTLHTYFNVNMTHTY 265

QY 230 KONGAPAEKLIYGFPTYGHNFILSNPNTGIGAPTSAGPAGPYAKESGIWYAYEIC-TF 288
DB 266 LNNATRDVLVWGVFPYGRWSTEDRSKULGDPKAGMSPPGFISGEEGVLSYIELCQLF 325

QY 289 LKNGATQGDAPQEVYAYQGNVWVGYDNIKSPDIKAQMLKHNKFGAMVWALDLDFTG 348
DB 326 QKEEMHIQYDEYVYAPYGNKDIWVGYDDLASTCKLAFELKELGVSGVWVWSLENDDFKG 385

QY 349 TFCNQKQKFLI-----STUKKALGLOSASCTAPAPDIEPITAAPSGS 390
DB 386 -HCGP-KNPLLNKVNHWINGDEKNSPECILGPTTPTPTPTPTPTPTPTPTPTPT 437

RESULT 14
US-09-923-844B-2
Sequence 2, Application US/09923844B
Patent No. US20020166143A1
GENERAL INFORMATION:
APPLICANT: Pioneer Hi-Bred International, Inc.
APPLICANT: Bao, Zhongmeng
APPLICANT: Lu, Guihua
TITLE OF INVENTION: Sclerotinia-inducible Genes and
TITLE OF INVENTION: Promoters and Their Uses
FILE REFERENCE: 35718/234631
CURRENT APPLICATION NUMBER: US/09/923,844B
CURRENT FILING DATE: 2001-08-07
PRIORITY APPLICATION NUMBER: US 60/224,603

PRIOR FILING DATE: 2000-08-11
NUMBER OF SEQ ID NOS: 20
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 371
TYPE: PRT
ORGANISM: Helianthus annuus
US-09-923-844B-2

Query Match 18.8%; Score 472; DB 9; Length 371;
Best Local Similarity 34.1%; Pred. No. 7.5e-32;
Matches 118; Conservative 47; Mismatches 145; Indels 36; Gaps 12;

QY 6 YETNWAQ-YRPGGRMPDNIDPCLCTHLIYAFAGRONNEITIEWN-----DVTLYQAFN 60
DB 30 YWPSWAQDFLP-----PSNIQTAYFTVYVYAFSLPNN---VTFQFDVHRTTASALNSFN 80

QY 61 -GLKKNKNSQLTLLAIGGNFCTAP-FTAMVSTPENROFTITSVIFLQRYEFDGLDFDM 118
DB 81 TALHGNKPNPKTLFSGGSGAGVKOLFSLKASSPGSRAAFIRSTIQVARNYYFDGADLDM 140

QY 119 EYPSGRSGPPQDKHLFTVLVQEMREAFQEAQKINPKRLMTAAVAAGISNIQSG---YE 175
DB 141 EYPTQ-----TDMNFGLLLDWVAVNNEATSTCKPRLLLSAATHEPEVDNGVAKYP 196

QY 176 IPQLSOYLDYIHVMYDYLHGSW-EGYTGNSPLYKYPTDTGSNAYLVNDVYMNWKNGA 234
DB 197 VASINKNLGDNAMCYDYGHPWTPDATGAPAAALY-----NPNGLSTSNGLQSWISAGI 250

QY 235 PAEKLVGFPTYGHNFILSNPNTGIGAPTSAGPAGPYAKESGIWYAYEICTF-LKNGA 293
DB 251 QROKLVGMGPLYGVTWKLKPNVNGIGAPAGGIGPG-----NEGAMLYSEVOOFNAQNNA 305

QY 294 TQGDWAPQEVYAYQGNVWVGYDNIKSPDIKAQMLKHNKFGAMVW 339
DB 306 RVYDTQTVSYISYSGTTWIGYDDVNSVQRYAKSLNIGGYFFW 351

RESULT 15
US-09-748-033-3
Sequence 3, Application US/09748033
Patent No. US20020069431A1
GENERAL INFORMATION:
APPLICANT: Broadway, Roxanne M.
APPLICANT: Gongora, Carmanza E.
TITLE OF INVENTION: EFFECT OF ENDOCHITINASE AND CHITOBIOCIDASE AND THEIR
TITLE OF INVENTION: ENCODING GENES ON PLANT GROWTH AND DEVELOPMENT
FILE REFERENCE: 19603/3091
CURRENT APPLICATION NUMBER: US/09/748,033
CURRENT FILING DATE: 2000-12-22
PRIORITY APPLICATION NUMBER: 60/172,003
PRIORITY FILING DATE: 1999-12-23
NUMBER OF SEQ ID NOS: 8
SOFTWARE: Patent in Ver. 2.1
SEQ ID NO 3
LENGTH: 376
TYPE: PRT
ORGANISM: Streptomyces albidoflavus
US-09-748-033-3

Query Match 11.4%; Score 285.5; DB 10; Length 376;
Best Local Similarity 28.0%; Pred. No. 4.7e-16;
Matches 117; Conservative 43; Mismatches 145; Indels 113; Gaps 25;

QY 6 YETNWAQYRPGGR-FMPDNI-----DPCLCTHLIYAFAGRONNEITIE-----49
DB 14 YFTWGVY-----GRNYHVNKLVLTSGSAEKITHINSGVNGGCKCTIGDSFAAYDKAYTA 69

QY 50 -----WNDVTLQAFN---GLKKNKNSQLTLLAIGGNFCTAPFTAMVSTPENRQ 96
DB 70 AESVDGVADTW-DQPLRGNFNLRLKAKYPHIKVLVSGGWTW-SGGFTDAVKPAAPA 127

QY 97 TFIITSVIFLQRYE-FDGLDFWMEYFGRS-----SPQDKHLFTVLVQEMREAFQEA 149

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2003, 20:56:54 ; Search time 14.1346 Seconds
(without alignments)
3094.613 Million cell updates/sec

Title: US-10-004-219b-14
Perfect score: 2506
Sequence: 1 YQLTCYFTNWAQYRPLGRF.....QQNCOAGLVFTSDCCNWA 455

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR73:***
1: pir1:***
2: pir2:***
3: pir3:***
4: pir4:***

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1361.5	54.3	399	2 S27879	secretory protein
2	976.5	39.0	383	2 A49562	cartilage glycopro
3	968.5	38.6	383	2 S51327	heparin-binding gl
4	959.5	38.3	654	2 I38605	oviductal glycopro
5	943.5	37.6	537	2 S57197	oviduct-specific g
6	941.5	37.6	539	2 I46470	estrogen dependent
7	912.5	36.4	483	2 A53918	chitinase (EC 3.2.
8	885	35.3	617	2 T15408	hypothetical prote
9	874	34.9	525	2 T44445	chitinase (EC 3.2.
10	863	34.4	554	2 A56596	chitinase (EC 3.2.
11	801.5	32.0	405	2 S61551	breast-regressing
12	786.5	31.4	504	2 A38221	chitinase (EC 3.2.
13	768.5	30.7	1635	2 T14075	chitinase (EC 3.2.
14	488.5	19.5	1215	2 T43916	chitinase A [import
15	482	19.2	599	2 D83764	chitinase BH0916 [
16	472.5	18.9	423	2 JQ1975	chitinase (EC 3.2.
17	465.5	18.6	398	2 T04761	chitinase (EC 3.2.
18	463.5	18.5	424	2 S47133	chitinase (EC 3.2.
19	463.5	18.5	699	2 A83368	chitinase (EC 3.2.
20	457.5	18.3	1484	2 T29275	hypothetical prote
21	456.5	18.2	756	2 AB1452	chitinase B homolo
22	454.5	18.1	2025	2 T03884	hypothetical prote
23	451.5	18.0	756	2 AB1088	chitinase B homolo
24	450	18.0	831	2 T00323	chitinase (EC 3.2.
25	444.5	17.7	379	2 T04762	chitinase homolog
26	440.5	17.6	424	2 S68121	chitinase I precur
27	434.5	17.3	423	2 S51369	chitinase - fungus
28	427.5	17.1	427	2 JC4565	chitinase (EC 3.2.
29	427	17.0	546	2 F84238	chitinase [importe

30	402	16.0	378	2 S51591	chitinase (EC 3.2.
31	393	15.7	366	2 T04763	chitinase homolog
32	388.5	15.5	511	2 S61166	probable membrane
33	388	15.5	452	2 JC4038	47K glycoprotein p
34	374.5	14.9	371	2 T04756	chitinase homolog
35	369.5	14.7	365	2 T04757	chitinase homolog
36	369.5	14.7	1051	2 D82428	chitodextrinase VC
37	368	14.7	633	2 T24898	hypothetical prote
38	367.5	14.7	869	2 T44440	chitinase (EC 3.2.
39	357.5	14.3	849	2 D82510	chitinase VCA0027
40	352.5	14.1	1054	2 T30933	chitinase (EC 3.2.
41	349.5	13.9	563	2 S06651	chitinase precursor
42	347.5	13.9	499	2 S52422	chitinase (EC 3.2.
43	345.5	13.8	332	2 T04754	hypothetical prote
44	344.5	13.7	499	2 S04856	chitinase (EC 3.2.
45	343	13.7	1046	2 T30199	chitinase (EC 3.2.

ALIGNMENTS

RESULT 1

S27879
secretory protein YM-1 precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 22-Jun-1999
C:Accession: S27879
R:Chang, N.C.A.; Liu, C.H.; Chang, A.C.
submitted to the EMBL Data Library, June 1992
A:Description: Molecular characterization of a secretory protein (YM-1) transiently expre
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-399/Product: secretory protein YM-1 #status predicted <MAT>
A:Reference number: S27879
A:Accession: S27879
A:Molecule type: mRNA
A:Residues: 1-399 <CHA>
A:Cross-references: EMBL:M94584; NID:G202441; PIDN:AAB62394.1; PID:G202442
C:Superfamily: Streptomyces chitinase chi40
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-399/Product: secretory protein YM-1 #status predicted <MAT>

Query Match 54.3%; Score 1361.5; DB 2; Length 399;
Best Local Similarity 66.3%; Pred. No. 7.3e-95; Indels 1; Gaps 1;
Matches 250; Conservative 46; Mismatches 80;

QY	1	YQLTCYFTNWAQYRPLGRFMPDNIDPCLCTHLIYAFAGRQNNETITIEWNDVLYQAFN	60
DB	22	YQLMCIYTSWAKDRPIEGSFKPGNIDPCLCTHLIYAFAGMNNETIYTHEQDLRDYEALN	81
QY	61	GLKN-KNSQLTKLLAIGGNFCTAPFTAMVSTPENRQFTITSVIKFLQYETDGLDFDWE	119
DB	82	GLKDKKNTLTKLLAIGGNFCTAPFTAMVSTPENRQFTITSVIKFLQYETDGLDFDWE	141
QY	120	YPSRSGSPQDKHLFTVLVOEMREAFQEAQINKPRLMTAAVAAGISNIOGVEIPQL	179
DB	142	YPSRSGSPQDKHLFTVLVOEMREAFQEAQINKPRLMTAAVAAGISNIOGVEIPQL	201
QY	180	SOYLDYHWYVDLHSGWEGYTGNSPLYKYPTDGTGSNAYLVNDVYVMYKNDGAPAEKL	239
DB	202	LSLLDYIQWYVDLHSGWEGYTGNSPLYKYPTDGTGSNAYLVNDVYVMYKNDGAPAEKL	261
QY	240	IVGFPTYGHNFILSNPSNTGICAPTSAGAGPAGYAKESGIWAYYEICTPLKNGATGWD	299
DB	262	IVGFPTYGHNFILSNPSNTGICAPTSAGAGPAGYAKESGIWAYYEICTPLKNGATGWD	321
QY	300	POEVPYAGQVWVGVYDNIKSFIDKAQMLKHNKFGAMVWADLDLDTCTCTFCNQGKFP	359
DB	322	POEVPYAGQVWVGVYDNIKSFIDKAQMLKHNKFGAMVWADLDLDTCTCTFCNQGKFP	381
QY	360	STLKKALGLQASCTAP	376
DB	382	STLKKALGLQASCTAP	398

RESULT 2

A49562
cartilage glycoprotein gp39 precursor - human
N;Alternate names: 39K synovial protein
C;Species: Homo sapiens (man)
C;Date: 23-Mar-1995 #sequence_revision 23-Mar-1995 #text_change 22-Jun-1999
C;Accession: A49562; S10677; A33162
R;Hakala, B.E.; White, C.; Recklies, A.D.
J. Biol. Chem. 268, 25803-25810, 1993
A;Title: Human cartilage gp-39, a major secretory product of articular chondrocytes and
A;Reference number: A49562; MUID:94064658; PMID:8245017
A;Accession: A49562
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-383 <HA>
A;Cross-references: GB:M80927; NID:G348911; PIDN:AAA16074.1; PID:G348912
R;Nyikos, P.; Golds, E.E.
Biochem. J. 269, 263-268, 1990
A;Title: Human synovial cells secrete a 39 kDa protein similar to a bovine mammary prote
A;Reference number: S10677; MUID:90328983; PMID:2375755
A;Accession: S10677
A;Molecule type: protein
A;Residues: 22-40, 'X', 42-45 <NY2>
C;Superfamily: Streptomyces chitinase chi40
C;Keywords: cartilage; extracellular protein; glycoprotein
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-383/Product: cartilage glycoprotein gp39 #status predicted <MAT>

Query Match 39.0%; Score 976.5; DB 2; Length 383;
Best Local Similarity 51.2%; Pred. No. 6.7e-66;
Matches 188; Conservative 62; Mismatches 108; Indels 9; Gaps 5;

QY 1 YOLTCTFTNWAQYRGLGRFMPDNIDPCLCTHLIYAFAGRQNNETITIEWNDVTLQAFN 60
DB 22 YKLVCTYTSWSQYREGDSCFPDLPFLCTHLIYAFANISNEIDTLEWNDVTLQAFN 81

QY 61 GLKNKNSQLTKLLAIGGNWFGTAPFTAMVSTPENRQTFTSVIKFLQYEFGLDFDWEY 120
DB 82 TLKKNPNLTKLLSVGGWNGFSQRFKTSASNTQSRRTIKSVPPFLRTHGFDGLDAMI 141

QY 121 PGRSGSPQDKHLFTVLVQEMREAFQEAQKQINRPLMTAAVAAGISNIQSGYEIPQLS 180
DB 142 PGRS-----DKQHFTTLIKEMKAEFTKEA-QPGKKQLLSAALSAGKVTIDISYDI 195

QY 181 QYLDYIHVMYDHLGWSWEGYTGNSPLKYPTDTGNSNAYLVNDVYVMYKNGAPAEKLI 240
DB 196 QHLDFTSIMTYDFHGNWGTGTHHSPLFRQGDASDRFSNADYAVSYVLRGLGAPAKLV 255

QY 241 VGPPTYGHNFLSNPNTGIGAPTSGAGPAGPYAKESGIWAYYEICTFLKNGATQGWDA 300
DB 256 MGIPTFGRSFTLAS-SETGVGAPISGPGIPGRFTKEAGTLAYEICDFLR-GATVHR 313

QY 301 QVPYAYQGNVWGYDNIKSFDIKAQWLKHNKFGGAMWALDLDFTGTFNCQG-KFPLI 359
DB 314 QQVPYATKGNQWGYDDQESVKSQYLRQDLGAMWALDLDFTGSGFCQGDLPFLT 373

QY 360 STLKAL 366
DB 374 NAIKDAL 380

RESULT 3
S51327
heparin-binding glycoprotein 38K - pig
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 22-Jun-1999
C;Accession: S51327
R;Shackleton, L.M.; Mann, D.M.; Millis, A.J.T.
submitted to the EMBL Data Library, January 1995
A;Description: Identification of a 38kDa heparin-binding glycoprotein (gp38k) in differe
A;Reference number: S51327
A;Accession: S51327
A;Status: preliminary
A;Molecule type: mRNA

A;Residues: 1-383 <SHA>
A;Cross-references: EMBL:247803; NID:G634097; PIDN:CAA87764.1; PID:G634098
C;Superfamily: Streptomyces chitinase chi40

Query Match 38.6%; Score 968.5; DB 2; Length 383;
Best Local Similarity 51.5%; Pred. No. 2.7e-65;
Matches 189; Conservative 57; Mismatches 112; Indels 9; Gaps 5;

QY 1 YOLTCTFTNWAQYRGLGRFMPDNIDPCLCTHLIYAFAGRQNNETITIEWNDVTLQAFN 60
DB 22 YKLVCTYTSWSQYREGDSCFPDLPFLCTHLIYAFANISNEIDTLEWNDVTLQAFN 81

QY 61 GLKNKNSQLTKLLAIGGNWFGTAPFTAMVSTPENRQTFTSVIKFLQYEFGLDFDWEY 120
DB 82 TLKKNPNLTKLLSVGGWNGFSQRFKTSASNTQSRRTIKSVPPFLRTHGFDGLDAMI 141

QY 121 PGRSGSPQDKHLFTVLVQEMREAFQEAQKQINRPLMTAAVAAGISNIQSGYEIPQLS 180
DB 142 PGRS-----DKRHFTTLVKEMKAEFVREALP-GTERLLLSGAVSAGKVAIDRGYDIAQ 195

QY 181 QYLDYIHVMYDHLGWSWEGYTGNSPLKYPTDTGNSNAYLVNDVYVMYKNGAPAEKLI 240
DB 196 QHLDFTSIMTYDFHGNWGTGTHHSPLFRQGDASDRFSNADYAVSYVLRGLGAPAKLV 255

QY 241 VGPPTYGHNFLSNPNTGIGAPTSGAGPAGPYAKESGIWAYYEICTFLKNGATQGWDA 300
DB 256 MGIPTFGRSFTLAS-SKTDVGAPASGPGIPGRFTKEAGTLAYEICDFLR-GATVRR 313

QY 301 QVPYAYQGNVWGYDNIKSFDIKAQWLKHNKFGGAMWALDLDFTGTFNCQG-KFPLI 359
DB 314 QQVPYATKGNQWGYDDQESVKNKAKYLKSLQLAGAMWMTLDLDDFRGNFCQNLRFPLT 373

QY 360 STLKAL 366
DB 374 NAIKDAL 380

RESULT 4
I38605
oviductal glycoprotein - human
C;Species: Homo sapiens (man)
C;Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 29-May-1998
C;Accession: I38605
R;Arias, E.B.; Verhage, H.G.; Jaffe, R.C.
Biol. Reprod. 51, 685-694, 1994
A;Title: Complementary deoxyribonucleic acid cloning and molecular characterization of a
A;Reference number: I38605; MUID:95119256; PMID:7819450
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-654 <RES>
A;Cross-references: EMBL:U09550; NID:G529147; PID:G529148

Query Match 38.3%; Score 959.5; DB 2; Length 654;
Best Local Similarity 49.3%; Pred. No. 2.6e-64;
Matches 187; Conservative 56; Mismatches 129; Indels 7; Gaps 3;

QY 1 YOLTCTFTNWAQYRGLGRFMPDNIDPCLCTHLIYAFAGRQNNETITIEWND-VTLQAF 59
DB 22 HKLVCTYTNWAHSRFGPASILPHDLPFLCTHLIYAFASMNQIVAKDQEKILYPEF 81

QY 60 NGLKNKNSQLTKLLAIGGNWFGTAPFTAMVSTPENRQTFTSVIKFLQYEFGLDFDWEY 119
DB 82 NKLKERNELTKLLSIGGNWNGFSFTTLMSTFANREKFIASVISLLRTHDGLDLPFL 141

QY 120 YPGRSGSPQDKHLFTVLVQEMREAFQEAQKQINRPLMTAAVAAGISNIQSGYEIPQL 179
DB 142 YPGLRGSPMDRWRTFLFLIEELLFAFRKEALLTMRPILLSAASGVPHIVQTSYDVRFL 201

QY 180 SQYLDYIHVMYDHLGWSWEGYTGNSPLKYPTDTGNSNAYLVNDVYVMYKNGAPAEKLI 239
DB 202 GRLLDFINVLSDYHLSWERTFGHNSPLFSFDPEDPKSSA-----YAMNYWRKLGAPSEKL 256

QY 240 IVGFPTYGHNFTLSPNTGICAPTSAGPAGYAKESGIWAYYEICTFLKNGATQGWDA 299
DB 257 IMGITYGRTFRLKAKSNGLOARAGPASPKYTKQSGFLAYFEICSPVW-GAKKHWD 315
QY 300 POEVPYAVOGNWWGVYDNKSPDIKAQMLKHNKFGAMVAIDLDDFTGTCNQGKFLPI 359
DB 316 YQYVYANKGKWEVGYDNAISFYKAWFIRREHFGAMVWTLDDMDVDRGTGCTGTPPFLV 375
QY 360 STLKKALGQSASCTAPAQ 378
DB 376 YVNDILVRAEFSSTSLPQ 394
RESULT 5
S57197
oviduct-specific glycoprotein 95K precursor - bovine (fragment)
C:Species: Bos primigenius taurus (cattle)
C:Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 05-Nov-1999
C:Accession: S57197
R:Sendai, Y.; Abe, H.; Kikuchi, M.; Satoh, T.; Hoshi, H.
Biol. Reprod. 50, 927-934, 1994
A:Title: Purification and molecular cloning of bovine oviduct-specific glycoprotein.
A:Reference number: S57197; MUID:94257768; PMID:8199272
A:Accession: S57197
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-537 <SEN>
A:Cross-references: EMBL:D16639; NID:g391621; PID:BAA04065.1; PID:d1004583; PID:g391622
C:Keywords: glycoprotein
Query Match 37.6%; Score 943.5; DB 2; Length 537;
Best Local Similarity 49.6%; Pred. No. 3.2e-63;
Matches 182; Conservative 53; Mismatches 125; Indels 7; Gaps 3;
QY 1 YOLTCYFTNWAQYRGLGRFMPDNIDPCLCTHLIYAFAGRONNEITTEWND-VTLIYQAF 59
DB 19 HKLVCFYTNWAFSRPGASILPRDLDPFLCTHLVFAFASMNQIVPKDPQDEKILYPEF 78
QY 60 NGLKKNKSLQTLAIGGNFCTAPFTAMVSTPENROTFTITSVIFKROYEDGLDFWE 119
DB 79 NKLKERNRGLKTLISGVGNFCTVFTMTLSFRERFVSVIALTRHTHGFGLDLFL 138
QY 120 YPGSRGSPQDKHLFTVLVQEMREAFEOAKQINPKRLMTVAAGAISNIQSQYEIPQL 179
DB 139 YPLGSGPARDRWTFVLELLQAFKNEAQTWRPRLLSAAVSGDHPVQKAYEARLL 198
QY 180 SOYLDYIHVMYDHLGSGWEGYTGNSPLYKYPTDTGSNAYLVNVDYVMYWKDNGAPAEKL 239
DB 199 GRLLDFISVLSYDLHGSWEKVTGHSPLSLPGDPKSSA-----YAMNWRQLGVPPPEKL 253
QY 240 IVGFPTYGHNFTLSPNTGICAPTSAGPAGYAKESGIWAYYEICTFLKNGATQGWDA 299
DB 254 LMGLPTYGRTFHLKASQNELRAQVGPASPKYTKQAGFLAYEICCVR- AKKRWIN 312
QY 300 POEVPYAVOGNWWGVYDNKSPDIKAQMLKHNKFGAMVAIDLDDFTGTCNQGKFLPI 359
DB 313 DQYVPYAFKGEWGYDDAISFGYKAFFIKREHFGAMVWTLDDDFRGYFGCTGTPPFLV 372
QY 360 STLKKAL 366
DB 373 HTLNLL 379
RESULT 6
I46470
estrogen dependent oviduct protein precursor - sheep
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C:Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 05-Nov-1999
C:Accession: I46470
R:DeSouza, M.M.; Murray, M.K.
Endocrinology 136, 2485-2496, 1995
A:Title: An estrogen-dependent secretory protein, which shares identity with chitinases,
and embryo development.

A:Reference number: I46470; MUID:95269691; PMID:7750470
A:Accession: I46470
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-539 <DES>
A:Cross-references: EMBL:U16719; NID:g885600; PIDN:AAC48471.1; PID:g885601
Query Match 37.6%; Score 941.5; DB 2; Length 539;
Best Local Similarity 48.8%; Pred. No. 4.6e-63;
Matches 179; Conservative 59; Mismatches 122; Indels 7; Gaps 3;
QY 1 YOLTCYFTNWAQYRGLGRFMPDNIDPCLCTHLIYAFAGRONNEITTE-WNDVTLIYQAF 59
DB 22 HKLVCFYTNWAFSRPGASILPRDLDPFLCTHLVFAFASMNQIVPKDPLEKILYPEF 81
QY 60 NGLKKNKSLQTLAIGGNFCTAPFTAMVSTPENROTFTITSVIFKROYEDGLDFWE 119
DB 82 NKLKERNRGLKTLISGVGNFCTVFTMTLSFRERFVSVIALTRHTHGFGLDLFL 141
QY 120 YPGSRGSPQDKHLFTVLVQEMREAFEOAKQINPKRLMTVAAGAISNIQSQYEIPQL 179
DB 142 YPLGSGPARDRWTFVLELLQAFKNEAQTWRPRLLSAAVSGDHPVQKAYEARLL 201
QY 180 SOYLDYIHVMYDHLGSGWEGYTGNSPLYKYPTDTGSNAYLVNVDYVMYWKDNGAPAEKL 239
DB 202 GRLLDFISVLSYDLHGSWEKVTGHSPLSLPGDPKSSA-----YAMNWRQLGVPPPEKL 256
QY 240 IVGFPTYGHNFTLSPNTGICAPTSAGPAGYAKESGIWAYYEICTFLKNGATQGWDA 299
DB 257 LMGLPTYGRTFHLKASQNELRAQVGPASPKYTKQAGFLAYEIVCSFVOR- AKKRWIN 315
QY 300 POEVPYAVOGNWWGVYDNKSPDIKAQMLKHNKFGAMVAIDLDDFTGTCNQGKFLPI 359
DB 316 DQYVPYAFKGEWGYDDAISFGYKAFFIKREHFGAMVWTLDDDFRGYFGCTGTPPFLV 375
QY 360 STLKKAL 366
DB 376 HTLNLL 382
RESULT 7
A53918
chitinase (EC 3.2.1.14) precursor - braconid wasp (Chelonus sp.)
C:Species: Chelonus sp.
C:Date: 28-Jul-1995 #sequence_revision 28-Jul-1995 #text_change 21-Jul-2000
C:Accession: A53918
R:Krishnan, A.; Nair, P.N.; Jones, D.
J. Biol. Chem. 269, 20971-20976, 1994
A:Title: Isolation, cloning, and characterization of new chitinase stored in active form
A:Reference number: A53918; MUID:94342256; PMID:8063715
A:Accession: A53918
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-483 <KRI>
A:Cross-references: GB:U10422; NID:g533504; PIDN:AAA61639.1; PID:g533505
C:Keywords: glycosidase; hydrolase; polysaccharide degradation
Query Match 36.4%; Score 912.5; DB 2; Length 483;
Best Local Similarity 38.0%; Pred. No. 6.1e-61;
Matches 185; Conservative 82; Mismatches 157; Indels 63; Gaps 11;
QY 2 QLTCTYFTNWAQYRGLGRFMPDNIDPCLCTHLIYAFAGRONNEITTE-WND---VTLIY 56
DB 23 KVCYFGAWSVTRQNGKFGDINGDPTLCTHLIYSGVGNCKDVKVDLPWSLDPLGNLDGF 82
QY 57 QAFNGLKKNKSLQTLAIGGNFCTAPFTAMVSTPENROTFTITSVIFKROYEDGLDF 116
DB 83 GKFTSLRKNKNSVKIMVAVGGMNAGSVFSSQWASDQATREAFQNVKFLQYQYQDFGDI 142
QY 117 DWEPYSGRSGSPQDKHLFTVLVQEMREAFEOAKQINPKRLMTVAAGAISNIQSQYEI 176
DB 143 DWEPYQARGGSPADYKNNVMVGLCKALKKAFVQH-----DYILSAVAAPETSASKSYDI 195

A56596
Chitinase (EC 3.2.1.14) - tobacco hornworm
C:Species: Manduca sexta (tobacco hornworm)
C>Date: 11-Aug-1995 #sequence_revision 11-Aug-1995 #text_change 29-Jan-1999
C:Accession: A56596
R:Kramer, K.J.; Corpuz, L.; Choi, H.K.; Muthukrishnan, S.
A:Title: Sequence of a cDNA and expression of the gene encoding epidermal and gut chitinase from Manduca sexta.
A:Reference number: A56596; MUID:93357793; PMID:8353525
A:Accession: A56596
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-554 <KRA>
A:Cross-references: GB:U02270; GB:S64757; NID:G406048; PID:G406049
A:Experimental source: larvae
A:Note: sequence extracted from NCBI backbone (NCBI:136417, NCBI:136418)
C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 34.4%; Score 863; DB 2; Length 554;
Best Local Similarity 35.1%; Pred. No. 3.9e-57;
Matches 185; Conservative 89; Mismatches 171; Indels 82; Gaps 13;

QY 2 QLTCTYTNWAOYRPGGLRPMFNDIDPCLTCLHLYAFAG--RQNNETITWNV---DVTLY 56
DB 24 RIVCYFSNVAVRPGVGRYIEDIPVEKCTHIYFISFIVGTGENSELVLIIDPELDVDKNGF 83
QY 57 QAFNGLKNKNSQLKTLAIGCNWFGTAPFTAMVSTPENRQTFITSVIKFLQYRFDGLDF 116
DB 84 RNFTSLRSHSPSVKPMVAVGWGAEGSSKYSHMVAQKSTRMSPFIRSVFSLKKYDFDGLDL 143
QY 117 DWEPGSG--RSGPPDQKHFLTVLQVEMREAFQEAQINKPRLMVTAAVAAGISNIQSGY 174
DB 144 DWEPGAGADRGSFSDKFLYLVLQELRAFRIRVCKGWE-----LTAAPVLANFLMEGY 198
QY 175 BIPQSLQYLDYIHWMTYDLHSGWEGYTGNSPLKYKPTDTGNSNAYLVNDYVMNWKONGA 234
DB 199 HVPCLQELDAIHWMSYDLRGNWAGFADVHSPLYKRPDQWYAEKLVNDGLHLWEKGC 258
QY 235 PAEKLVGFPYGHNFILS--NPSNTGIGA---PTSGAGPAGPYAKESGIWAYYEICTFL- 289
DB 259 FSNKLVGPIFYGRSFTLSAGNNYGLGTFINKEAGGDPAPYTNATGFWAYYEICTEVD 318
QY 290 --KNGATCGWDAPOBPVYACQNVWVDNKSFDIKAQWLKHNKFGAMVWAIIDDDFT 347
DB 319 KDDSWTKWDEQKCPYAKYGTQWGVGDEPRSEIKMNIKQKGYLGAMTWAIIDDDFO 378
QY 348 GTFCNQKFPFLITLKKALQASAS-----CTAPAOPIEP-----ITA 385
DB 379 G-LCGE-KNPLIKILHKEMSSVTPPPHTENTPTPEWARPPSTPSPSEGDPIPTTTA 436
QY 386 APSGSGNSGSSSSGSS-----GGSG 407
DB 437 KPASTTKTVKTTTAKPPQSVIDEENDINVRPEKPEPEPEPEVEVPPTENEVDGSE 496
QY 408 FCVAVRANGLYPVANNRNFAHCVGVTVQNCQAGLVFDTSCDCNWW 454
DB 497 ICNSQDII-PDKKCHDKYRCVNGEAMQFSCQHGTVFNVELNVCDW 542

RESULT 11
S61551.
breast-regressing protein brp39 precursor - mouse
C:Species: Mus musculus (house mouse)
C>Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
C:Accession: S61551; S61550; I48271
R:Morrisson, B.W.; Leder, P.
A:Title: new and ras initiate murine mammary tumors that share genetic markers generally
A:Reference number: I48271; MUID:195060797; PMID:7970700
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-405 <MORI>

A:Cross-references: EMBL:X93035; NID:G1085065; PIDN:CRA63603.1; PID:G1085066
R:Morrisson, B.W.
Submitted to the EMBL Data Library, November 1995
A:Reference number: S61550
A:Accession: S61550
A:Molecule type: mRNA
A:Residues: 1-245, 'I', 247-330, 'H', 332-350, 'MVALDLDQFQCTQPKFFPLTNAIKDALA' <MOR2>
A:Cross-references: EMBL:X93035; NID:G1085065; PIDN:CRA63603.1; PID:G1085066
A:Note: the differences at the carboxyl end are due to a frameshift error
C:Genetics:
A:Gene: brp39
C:Superfamily: Streptomyces chitinase chi40
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-405/Product: breast-regressing protein brp39 #status predicted <MAT>

Query Match 32.0%; Score 801.5; DB 2; Length 405;
Best Local Similarity 45.7%; Pred. No. 1.1e-52;
Matches 158; Conservative 63; Mismatches 112; Indels 13; Gaps 6;

QY 1 YOLTCTYTNWAOYRPGGLRPMFNDIDPCLTCLHLYAFAG--RQNNETITWNVDTLYQAF 59
DB 22 YKLVCTYTSQYREGVGSFLPDAIQPLCTHIYSPANISSDNMLSTWENDESNDYKL 81
QY 60 NGLKNKNSQLKTLAIGCNWFGTAPFTAMVSTPENRQTFITSVIKFLQYRFDGLDFWE 119
DB 82 NKLKTRNTNLKTLSSVGWKEGKRFSEIASNTERRTAFVRSVAPFLRSYGFGLDLAWL 141
QY 120 YPSGRSGPPQDKHFLTVLQVEMREAFQEAQINKPRLMVTAAVAAGISNTQSGVEIQPL 179
DB 142 YPLRL-----DKQYFESTLIKELNAEFTKEV--QPGREKLLLSAALSAGKVAIDTGYDIAQI 195
QY 180 SOYLDYIHWMTYDLHSGWEGYTGNSPLKYKPTDTGNSNAYLVNDYVMNWKONGAPAEKL 239
DB 196 AQHLDYFNLMTYDFHGWVWRTTGHHSPLFQQCKQTRPDYRNVYAVQYMTLGAQAQSL 255
QY 240 IVGPFTYGHNFILSNPTGIGAPTSGAGPAGPYAKESGIWAYYEICTFLKNGATQGWDA 299
DB 256 LMGIPTFGKSTFLASSENO--LGAPISGEGLPGRFTKEAGTLAYYEICDFLK--GAEVHRLS 313
QY 300 PQEVPYAYQGNVWVDNKSFDIKAQWLKHNKFGGA-----MVWAI 341
DB 314 NEKVPFATKGNQWGYEDKESVQNKVGFLEKELAGACGHWIWI 359

RESULT 12
A38221
chitinase (EC 3.2.1.14) MFI - nematode (Brugia malayi)
C:Species: Brugia malayi
C>Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jul-2000
C:Accession: A38221
R:Fuhrman, J.A.; Lane, W.S.; Smith, R.F.; Piessens, W.F.; Perlier, F.B.
Proc. Natl. Acad. Sci. U.S.A. 89, 1548-1552, 1992
A:Title: Transmission-blocking antibodies recognize microfilarial chitinase in brugian l
A:Reference number: A38221; MUID:92179220; PMID:1542646
A:Accession: A38221
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid; protein
A:Residues: 1-504 <FUH>
A:Cross-references: GB:M73689; NID:G156063; PIDN:AAA27854.1; PID:G156064
A:Note: sequence extracted from NCBI backbone (NCBI:85345)
C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 31.4%; Score 786.5; DB 2; Length 504;
Best Local Similarity 35.0%; Pred. No. 2e-51;
Matches 169; Conservative 73; Mismatches 200; Indels 41; Gaps 9;

QY 1 YOLTCTYTNWAOYRPGGLRPMFNDIDPCLTCLHLYAFAG--RQNNETITWNVDT----- 54
DB 23 YRGCYTNWAOYRDEGKFLPQNLGCTHIYAFPAKVDGLGSKPFEWDETEWSK 82
QY 55 -LYQAFNGLKNKNSQLKTLAIGCNWFGTAPFTAMVSTPENRQTFITSVIKFLQYRFDG 113
DB 83 GMSAVTKLRETNPLKLVLLSYGVYGFSAIFTGIKSAQKTERFIKSAIAFLRNKDFG 142

C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: D83764
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Maqui, N.; Fuji, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; NUID:20512582; PMID:11058132
A:Accession: D83764
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-599 <STO>
A:Cross-references: GB:AP001510; GB:BA000004; NID:g10173440; PIDN:BA04635.1; GSPDB:GN00
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH916

	Query Match	19.2%;	Score 482;	DB 2;	Length 599;
	Best Local Similarity	27.9%;	Pred. No. 2.1e-28;		
	Matches 127;	Conservative 85;	Mismatches 144;	Indels 100;	Gaps 20;
Qy	1	YQLTCYTNWAOYRPLQGR-FMPDNIDPCLCTHLIYAF-----GRQNN-----	43		
Db	40	YKIVAYYFSWGAY-----GRDQVWDIDASKISHINYAFANICWDGRHGNPDGPNQWTW	95		
Qy	44	-----BITTIE-----WNVTL-----YQAFNGLKNKNSQLKTLA	74		
Db	96	SCQDENGVIDVPNGSIVMGDPWIDAKSNPGDTWDEPLRGNFQQLNKLKEEHPHLKTLIS	155		
Qy	75	IGGNWFGTAPTAMYSTPENQTEITSVKELROYEFDGLDPEWYVGRSGSP-----PO	129		
Db	156	VGGWTWSNR-FSDMAATKETRENFANSVEFIRKYGFDGVDVDWEYVPSGGLPGNSRRPE	214		
Qy	130	DKHLFTVLVQEMREAFQEQAKINKPRLMVTAAVAAGISNIQSGY-----EIPQLSQYLDY	185		
Db	215	DKENHVLLQEVRLKDEAGQEDGXDLL---TIASGAS---PGIVENNKLNEIAEIVDW	268		
Qy	186	IHMVTYDLHGSWEGVTGENSPLYKYP-----TDTGSNAVLYNDVYVMYKNDGAPAEKLIV	241		
Db	269	INIMTYDFNGGQWQISGHNAPLYDPATANTELPTPEHFNVESAVEGHLQAGVPEHKLVL	328		
Qy	242	GFPTYGHNFILSNPNTG-----IGAPTSCAGPAGPYAKESGIWAYYEICTFL--KNGATQ	295		
Db	329	GMPFYGRGWSNCDGANQGEYORCAPPREGTW-----ENGVFDFSLDEHYINKNGYOR	381		
Qy	296	GWDAPQEVPIYAY---QGNVWGVYDNIKSFIDIKAOMLKHNKFGGAMWALDLDFTGTFCN	352		
Db	382	YNDVAKVPFFLYNATNGN-FITYDDEESFRYKTDYFKSNLAGSNFWDVS-GDRNGT---	436		
Qy	353	QCKFPLISTLKKALGLQSASCTAPAPQIEPIAAPS	388		
Db	437	-----LLTALADQLGFTPEHQEPEEP-----SSAPT	463		

Search completed: June 29, 2003, 21:02:42
Job time : 15.1346 secs

THIS PAGE BLANK (0870)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run On: June 29, 2003, 20:55:24 ; Search time 25.5769 Seconds
(without alignments)
3665.470 Million cell updates/sec

Title: US-10-004-219b-14

Perfect score: 2506

Sequence: 1 YQLTCVFTNWAQYRGLGRF.....QONCOAGLVFTSDCCNWA 455

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL.21.*

- 1: sp archaea.*
- 2: sp bacteria.*
- 3: sp fungi.*
- 4: sp human.*
- 5: sp invertebrate.*
- 6: sp mammal.*
- 7: sp mhc.*
- 8: sp organelle.*
- 9: sp phase.*
- 10: sp plant.*
- 11: sp rodent.*
- 12: sp virus.*
- 13: sp vertebrate.*
- 14: sp unclassified.*
- 15: sp virus.*
- 16: sp bacteriap.*
- 17: sp archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	2506	100.0	476 4	Q9BZP6
2	2097.5	83.7	472 11	Q9JLN1
3	2097.5	83.7	473 11	Q9D803
4	2089.5	83.4	473 11	Q99PH2
5	2039	81.4	472 6	Q95M17
6	2010	80.2	368 4	Q9ULY4
7	1725	68.8	315 4	Q9ULY3
8	1670.5	66.7	365 11	Q9IXA9
9	1415	56.5	398 11	O35744
10	1407	56.1	398 11	P70201
11	1387	55.3	398 11	Q9IZ98
12	1384	55.2	402 11	O8VHG1
13	1379	55.0	402 11	O8VH43
14	1340.5	53.5	488 13	Q9OW34
15	1323	52.8	466 4	Q13231
16	1239	49.4	262 11	Q9D7W6

17	1173	46.8	387 4	Q9H3V8
18	1008.5	40.2	696 5	Q9VZV2
19	980	39.1	1013 5	Q96OM0
20	974.5	38.9	383 4	Q96HL7
21	972.5	38.8	383 6	Q8SPQ0
22	968.5	38.6	383 6	Q29411
23	966.5	38.6	390 4	Q96F97
24	966	38.5	260 11	O61201
25	954.5	38.1	624 6	O19118
26	948.5	37.8	352 11	Q9WTV1
27	933	37.2	457 6	Q95LB3
28	919.5	36.7	396 11	Q9D7Q1
29	912.5	36.4	483 5	Q23737
30	890.5	35.5	4498 5	Q9W223
31	881	35.2	332 6	O18949
32	879.5	35.1	381 11	O99J84
33	874	34.9	525 5	O44079
34	861	34.4	462 5	Q8SYH0
35	855	34.1	544 5	Q9GQC4
36	853	34.0	460 5	Q9W2M7
37	853	34.0	566 5	Q8WR52
38	852	34.0	543 5	Q9GV05
39	852	34.0	565 5	P90710
40	851	34.0	543 5	Q9GR93
41	849.5	33.9	553 5	P91731
42	848	33.8	552 5	Q9GV44
43	832	33.2	565 5	Q9PG9
44	828	33.0	520 5	O17100
45	822	32.8	524 5	Q27919

ALIGNMENTS

RESULT 1

Q9BZP6 ID Q9BZP6 PRELIMINARY; PRT; 476 AA.
 AC Q9BZP6; TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DE Acidic mammalian chitinase precursor (EC 3.2.1.14).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RC SEQUENCE FROM N.A.
 RC TISSUE=STOMACH, AND LUNG;
 RX MEDLINE=21125893; PubMed=11085997;
 RA Boot R.G., Blommaert E.F.C., Swart E., Ghauharali-Van Der Vlugt K.,
 RA Bijl N., Moe C., Place A., Aerts J.M.F.G.;
 RT "Identification of a Novel Acidic Mammalian Chitinase Distinct from
 RT Chitotriosidase.";
 RL J. Biol. Chem. 276:6770-6778 (2001).
 DR EMBL; AF290004; AAG60019.1;
 DR InterPro; IPR001579; Chitinase 18/2.
 DR InterPro; IPR002557; Chitin bind Per.
 DR InterPro; IPR001223; Glyco_hydro_18.
 DR Pfam; PF01607; CBM_14; 1.
 DR Pfam; PF00704; Glyco_hydro_18; 1.
 DR ProDom; PD000471; Glyco_hydro_18; 1.
 DR SMART; SM00494; ChitBD2; 1.
 DR PROSITE; PS01095; CHITINASE_18; UNKNOWN_1.
 KW Glycosidase; Hydrolase; Signal.
 FT SIGNAL
 SQ SEQUENCE 476 AA; 52271 NW; 92B27BAD2F7EB4CC CRC64;

Query Match 100.0%; Score 2506; DB 4; Length 476;

Best Local Similarity 100.0%; Pred. No. 1.5e-171; Indels 0;

Matches 455; Conservative 0; Mismatches 0; Gaps 0;

QY 1 YQLTCVFTNWAQYRGLGRFMPDNIDPCLCTHLIYAFAGRQNNETTIEWNDVTLTYQAFN 60


```

DR MGD; MGI:1932052; Chia.
DR InterPro; IPR001579; Chitinase_18/2.
DR InterPro; IPR002557; Chitin bind PerA.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF01607; CBM_14; 1.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR SMART; SM00494; ChtBD2; 1.
DR PROSITE; PS01095; CHITINASE_18; 1.
KW Glycosidase; Hydrolase.
SQ SEQUENCE 473 AA; 52003 MW; 333C874477476695 CRC64;

Query Match      83.7%; Score 2097.5; DB 11; Length 473;
Best Local Similarity 81.5%; Pred. No. 2.8e-142;
Matches 370; Conservative 40; Mismatches 41; Indels 3; Gaps 1;

QY 1 YQLTCYFTNWAQYRPGLGFMFNDIPCLCTHLIYAFAGQNNETITIEWNDVTLYQAFN 60
DB 22 YNLICYFTNWAQYRPGLGSKFPDDINPCLCTHLIYAFAGQNNETITIEWNDVTLYKAFN 81
QY 61 GLKKNNSQLKTLAIGGNFNGTAPFTAMVSTPENRQTFTITSVIKFLQYFEGDGLDWEY 120
DB 82 DLKRNNSKLKTLAIGGNFNGTAPFTTWVSTSONRQTFTITSVIKFLQYFEGDGLDWEY 141
QY 121 PGSRSGPPQDKHLFTVLVOEMREAEQEAQKINRPLMTAAVAGISNIQSGYIPLQS 180
DB 142 PGSRSGPPQDKHLFTVLVKEMREAEQEAIESNRPLMTAAVAGISNIQAGYIPELS 201
QY 181 QYLDVIHWTYDLHGSWEGYTGENSEPLYKYPTDTSNAYLVNDYVYNNKNGAPAEKLI 240
DB 202 KYLDVIHWTYDLHGSWEGYTGENSEPLYKYPTETGSNAYLVNDYVYNNKNGAPAEKLI 261
QY 241 VGFPYGHNFILSPNSNTGIGAPTSGAGPAGYAKESGIWAYEICTFLKNGATQGWADP 300
DB 262 VGFPYGHNFILSPNSDNGIGAPTSQDGPAGYTRQAGFWAYEICTFLRSGATEWADAS 321
QY 301 QEVPIYAGQNVWVYDNIKSFDIKAQWLKHNKFGGAMVAIDLDDPTGTFNCGKFPPLIS 360
DB 322 QEVPIYAYKANWLGVDNIKSFVSKAQLKQNNFGGAMIWAIDLDDPTGTFNCGKFPPLIS 381
QY 361 TLKKGALGLOSASCTAPAOPIETAPSGSGNGSGSSSGSGSGSGGFCACVANGLYPVA 420
DB 382 TLNKGALISTEGCTAPDVPSEPVTPPP---GSGSGSGSGSGSGSGSGGFCADKADGLYPVA 438
QY 421 NNRNAFWHCVNGVTYQONCOAGLVFDTSCDCNNW 454
DB 439 DDRNFAWQCINGITTYQOHCQAGLVFDTSCNCCNW 472

RESULT 4
Q99PH2 PRELIMINARY; PRT; 473 AA.
AC Q99PH2;
DT 01-JUN-2001 (TremBLrel. 17, Created)
DT 01-JUN-2001 (TremBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
DE Acidic mammalian chitinase (EC 3.2.1.14).
GN CHIA.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C; TISSUE=LUNG;
RX MEDLINE=21125893; PubMed=11085997;
RA Boot R.G., Blommaert E.F.C., Swart E., Ghauharali-Van Der Vlugt K.,
RA Bijl N., Moe C., Place A., Aerts J.M.F.G.;
RT Identification of a Novel Acidic Mammalian Chitinase Distinct from
RT Chitotriosidase.
RL J. Biol. Chem. 276:6770-6778(2001).
CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
HYDROLASES).

DR MGD; MGI:1932052; Chia.
DR InterPro; IPR001579; Chitinase_18/2.
DR InterPro; IPR002557; Chitin bind PerA.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF01607; CBM_14; 1.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR SMART; SM00494; ChtBD2; 1.
DR PROSITE; PS01095; CHITINASE_18; 1.
KW Glycosidase; Hydrolase.
SQ SEQUENCE 473 AA; 51977 MW; 389D87557BEC0784 CRC64;

Query Match      83.4%; Score 2089.5; DB 11; Length 473;
Best Local Similarity 81.3%; Pred. No. 1.1e-141;
Matches 369; Conservative 40; Mismatches 42; Indels 3; Gaps 1;

QY 1 YQLTCYFTNWAQYRPGLGFMFNDIPCLCTHLIYAFAGQNNETITIEWNDVTLYQAFN 60
DB 22 YNLICYFTNWAQYRPGLGSKFPDDINPCLCTHLIYAFAGQNNETITIEWNDVTLYKAFN 81
QY 61 GLKKNNSQLKTLAIGGNFNGTAPFTAMVSTPENRQTFTITSVIKFLQYFEGDGLDWEY 120
DB 82 DLKRNNSKLKTLAIGGNFNGTAPFTTWVSTSONRQTFTITSVIKFLQYFEGDGLDWEY 141
QY 121 PGSRSGPPQDKHLFTVLVOEMREAEQEAQKINRPLMTAAVAGISNIQSGYIPLQS 180
DB 142 PGSRSGPPQDKHLFTVLVKEMREAEQEAIESNRPLMTAAVAGISNIQAGYIPELS 201
QY 181 QYLDVIHWTYDLHGSWEGYTGENSEPLYKYPTDTSNAYLVNDYVYNNKNGAPAEKLI 240
DB 202 KYLDVIHWTYDLHGSWEGYTGENSEPLYKYPTETGSNAYLVNDYVYNNKNGAPAEKLI 261
QY 241 VGFPYGHNFILSPNSNTGIGAPTSGAGPAGYAKESGIWAYEICTFLKNGATQGWADP 300
DB 262 VGFPYGHNFILSPNSDNGIGAPTSQDGPAGYTRQAGFWAYEICTFLRSGATEWADAS 321
QY 301 QEVPIYAGQNVWVYDNIKSFDIKAQWLKHNKFGGAMVAIDLDDPTGTFNCGKFPPLIS 360
DB 322 QEVPIYAYKANWLGVDNIKSFVSKAQLKQNNFGGAMIWAIDLDDPTGTFNCGKFPPLIS 381
QY 361 TLKKGALGLOSASCTAPAOPIETAPSGSGNGSGSSSGSGSGSGGFCACVANGLYPVA 420
DB 382 TLNKGALISTEGCTAPDVPSEPVTPPP---GSGSGSGSGSGSGSGSGGFCADKADGLYPVA 438
QY 421 NNRNAFWHCVNGVTYQONCOAGLVFDTSCDCNNW 454
DB 439 DDRNFAWQCINGITTYQOHCQAGLVFDTSCNCCNW 472

RESULT 5
Q95M17 PRELIMINARY; PRT; 472 AA.
AC Q95M17;
DT 01-DEC-2001 (TremBLrel. 19, Created)
DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
DE Chitin binding protein b04.
GN CBP B04.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE=21475601; PubMed=11591385;
RA Suzuki M., Morimatsu M., Yamashita T., Iwanaga T., Syuto B.;
RT "A novel serum chitinase that is expressed in bovine liver.";
RL FEBS Lett. 506:127-130(2001).
CC EMBL; AB051629; BAB71805.1; -.
DR InterPro; IPR001579; Chitinase_18/2.

```

4.

```
QY 201 TGNSPLYKYPTDTGSGNAYLVNDYVYNNWKDNGAPAEKLIIVGPTVGHNFILSNPNTGI 260
Db 61 TGNSPLYKYPTDTGSGNAYLVNDYVYNNWKDNGAPAEKLIIVGPTVGHNFILSNPNTGI 120
QY 261 GAPTSGAGPAGYAKESGWAYEICTFLKNGATQGDAPQVVPYQGNVWGVYDNIS 320
Db 121 GAPTSGAGPAGYAKESGWAYEICTFLKNGATQGDAPQVVPYQGNVWGVYDNIS 180
QY 321 FDIKAQWLKHNKFGGAMVWAIDLDPTGTFNCGKFPPLISTLKALGLQSASCTAPAOPI 380
Db 181 FDIKAQWLKHNKFGGAMVWAIDLDPTGTFNCGKFPPLISTLKALGLQSASCTAPAOPI 240
QY 381 EBITAAPSGSGSGSGSGSGSGSGGFCVAVRANGLYPVANNRNPFWHCVNGVTVQOQCQ 440
Db 241 EBITAAPSGSGSGSGSGSGSGSGGFCVAVRANGLYPVANNRNPFWHCVNGVTVQOQCQ 300
QY 441 AGLVFDTCDCCNWA 455
Db 301 AGLVFDTCDCCNWA 315

RESULT 8
Q91XA9 PRELIMINARY; PRT; 365 AA.
AC Q91XA9;
DT 01-DEC-2001 (TREMELrel. 19, Created)
DT 01-DEC-2001 (TREMELrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMELrel. 21, Last annotation update)
DE Similar to eosinophil chemotactic cytokine.
GN CHIA.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC011134; AAH11134.1; -.
DR MGD; MGI:1932052; Chia.
DR InterPro; IPR001579; Chitinase_18/2.
DR InterPro; IPR002557; Chitin bind_Pera.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF01607; CBM 14; 1.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR PROSITE; PS01095; CHITINASE_18; UNKNOWN.1.
SQ SEQUENCE 365 AA; 39968 MW; C900BE0C4416F1DD CRC64;

Query Match 66.7%; Score 1670.5; DB 11; Length 365;
Best Local Similarity 80.4%; Pred. No. 7.9e-112;
Matches 295; Conservative 35; Mismatches 34; Indels 3; Gaps 1;

QY 88 MVSTPENRQTFTSVIKFLRQYEFGLDFWBYPSGRSPDQKHLFTVLVQEMREAFEQ 147
Db 1 MVSTSONRQTFTSVIKFLRQYGFGLDLDFWBYPSGRSPDQKHLFTVLVQEMREAFEQ 60
QY 148 BAKQINKPLMTAAVAGISINISQSYEIPQLSYLDYIHVMTYDLHGSWEGYTGENSPL 207
Db 61 EATGSNRPRLMTAAVAGGISINISQSYEIPQLSYLDYIHVMTYDLHGSWEGYTGENSPL 120
QY 208 KYPTDTGSNAYLVNDYVYNNWKDNGAPAEKLIIVGPTVGHNFILSNPNTGIGAPTSQA 267
Db 121 KYPTDTGSNAYLVNDYVYNNWKDNGAPAEKLIIVGPTVGHNFILSNPNTGIGAPTSQA 180
QY 268 GPAGYAKESGWAYEICTFLKNGATQGDAPQVVPYQGNVWGVYDNISFQDKAQW 327
Db 181 GPAGPYTRQAGWAYEICTFLRSGATEVWDASQVVPYKANEWLYGNDYNIKSFSVKAQW 240
QY 328 LKHNKFGGAMVWAIDLDPTGTFNCGKFPPLISTLKALGLQSASCTAPAOPIEPIITAP 387
Db 241 LKQNNFGGAMVWAIDLDPTGTFNCGKFPPLISTLKALGLQISTEGCTAPDVPFSEPTTTP 300
```

```
QY 388 SSGSGSGSGSGSGSGSGSGGFCVAVRANGLYPVANNRNPFWHCVNGVTVQOQCAGLVFD 447
Db 301 ---GSGSGSGSGSGSGSGSGGFCADKADGLYPVADDRNAPWOCINGITTYQHCQAGLVFD 357
QY 448 SCDCCNW 454
Db 358 SCNCCNW 364

RESULT 9
Q35744 PRELIMINARY; PRT; 398 AA.
AC Q35744;
DT 01-JAN-1998 (TREMELrel. 05, Created)
DT 01-MAR-2001 (TREMELrel. 16, Last sequence update)
DT 01-MAR-2002 (TREMELrel. 20, Last annotation update)
DE Secretary protein precursor.
GN CHI3L3 OR YMI.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Chang N.C.A.; Hung S.I.; Hwa K.Y.; Kato I.; Chen J.E.; Liu C.H.;
RA Chang A.C.;
RT "A macrophage protein, Ymi, transiently expressed during inflammation
RT is a novel mammalian lectin.";
RL J. Biol. Chem. 276:17497-17506 (2001).
DR EMBL; M94584; AAB62394.2; -.
DR MGD; MGI:1330860; Chi3l3.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
KW Signal.
FT SIGNAL.
FT CHAIN.
SQ SEQUENCE 398 AA; 44458 MW; C1187661B99D1D1 CRC64;

Query Match 56.5%; Score 1415; DB 11; Length 398;
Best Local Similarity 68.1%; Pred. No. 1.8e-93;
Matches 256; Conservative 48; Mismatches 72; Indels 0; Gaps 0;

QY 1 YOLTCYFTNWAQYRGLGREMPDNIDPCLCTHLIYAFAGRONNEITTIEMNDVTLYQAPN 60
Db 22 YQLMCTYTSWAKDRPIEGSFKPGNIDPCLCTHLIYAFAGMONEITYTHEQDLRDYEALN 81
QY 61 GLKKNNSQLKTLTLLAIGGNWFGTAPFTAMVSTPENRQTFTSVIKFLRQYEFGLDFWBY 120
Db 82 GLKKNNSQLKTLTLLAIGGNWFGTAPFTAMVSTPENRQTFTSVIKFLRQYEFGLDFWBY 141
QY 121 PGRSGSPDQKHLFTVLVQEMREAFQBAKQINKPLMTAAVAGISINISQSYEIPQLS 180
Db 142 PGRSGSPDQKHLFTVLVQEMREAFQBAKQINKPLMTAAVAGISINISQSYEIPQLS 201
QY 181 QYLDYIHVMTYDLHGSWEGYTGENSPLYKYPTDTGSGNAYLVNDYVYNNWKDNGAPAEKLI 240
Db 202 QSLDYIQVMTYDLHDKPDGQYTGNSPLYKSPYDYGKSDALNVDISIYWKHOGAASEKLI 261
QY 241 VGFPYTHNFILSNPNTGIGAPTSAGPAGYAKESGWAYEICTFLKNGATQGDAP 300
Db 262 VGFPYTHNFILSNPNTGIGAPTSAGPAGYAKESGWAYEICTFLKNGATQGDAP 321
QY 301 QEVPIYQGNVWGVYDNISFQDKAQINKPLMTAAVAGISINISQSYEIPQLS 360
Db 322 QEVPIYQGNVWGVYDNISFQDKAQINKPLMTAAVAGISINISQSYEIPQLS 381
QY 361 TLKALGLQSASCTAP 376
Db 382 TLKGLDNIHSASCRGP 397
```

```
RESULT 10
P70201
ID P70201 PRELIMINARY; PRT; 398 AA.
AC P70201;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-MAR-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2002 (TrEMBLrel. 20, Last annotation update)
DE ECF-L.
GN CHI3L3 OR MECF-L.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BONE MARROW;
RX MEDLINE=2009202; PubMed=10625674;
RA Ohashi M., Arita H., Hayai N.;
RT "Identification of a novel eosinophil chemotactic cytokine (ECF-L) as
RT a chitinase family protein.";
RL J. Biol. Chem. 275:1279-1286(2000).
DR EMBL; D87757; BAA13458.2; -
DR MGD; MG1:1330860; Chi3l3.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR Signal.
SQ SEQUENCE 398 AA; 4448 MW; 9D133FC490096793 CRC64;

Query Match 56.1%; Score 1407; DB 11; Length 398;
Best Local Similarity 67.8%; Pred. No. 6.8e-93;
Matches 255; Conservative 48; Mismatches 73; Indels 0; Gaps 0;

QY 1 YOLTCYFTNWAQYRPGLRMPNDIDPCLCTHLIYAFAGQNNNEITTIENWDTLYQAFN 60
DB 22 YOLMCYFTSWAKDRPTEGSKFGNIDPCLCTHLIYAFAGQNNNEITTIENWDTLYQAFN 81
QY 61 GLKKNKSOLKTLAIGGNWFGTAPFTAMYSTPNRQTFITSVIKELROYEPDGLDFDWEY 120
DB 82 GLKKNKTELKTLAIGGNWFGTAPFTAMYSTPNRQTFITSVIKELROYEPDGLDFDWEY 141
QY 121 PGSRGSPPQDKHLFTVLVQEMREAFQEAQKQINPKRLMVTAAVAGISNIQSGYIPLS 180
DB 142 PGSRGSPPKDKHLFSLVQEMKAFEEESTLNHI PRLLLTSTGAGFIDVIKSGYKIPLS 201
QY 181 QVLDIHWYTDLHGSWEGYTGENSEPLKYKPTDTSNAYLVNDYVNTWKONGAPAEKLI 240
DB 202 QSLDYIQVMTYDLHDPKDGTYGENSEPLKSPYDYGKADLNVDISITTYWKDHGAASEKLI 261
QY 241 VGFPTYGHNFIILSNPNTGIGAPTSGAGPAGPYAKESGIWAYVEICTFLKNGATQGDAP 300
DB 262 VGFPAHYHTFILDSPKGTIGAPTSTGPPGKYTDESGLLAYEVCITFLNEGTEWDAP 321
QY 301 QEVPIYQGNVWGYDNIKSFDIKAQWLKHNKFGGAMVWALDDFTGTFCNQKFPPLIS 360
DB 322 QEVPIYQGNVWGYDNIKSFDIKAQWLKHNKFGGAMVWALDDFTGTFCNQKFPPLIS 381
QY 361 TLKKAIGLOSASTAP 376
DB 382 TLKGLDNIHSASCKGP 397

RESULT 11
Q91298
ID Q91298 PRELIMINARY; PRT; 398 AA.
AC Q91298;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Secreted protein precursor Ym2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/OLA;
RX Su W.B., Chang N.-C.A.;
RT "Genomic organization of the ym2 gene.";
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF461142; AAL66748.1; -
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF00704; Glyco_hydro_18; 1.

Query Match 55.3%; Score 1387; DB 11; Length 398;
Best Local Similarity 66.5%; Pred. No. 1.8e-91;
Matches 250; Conservative 52; Mismatches 74; Indels 0; Gaps 0;

QY 1 YOLTCYFTNWAQYRPGLRMPNDIDPCLCTHLIYAFAGQNNNEITTIENWDTLYQAFN 60
DB 22 YOLMCYFTSWAKDRPTEGSKFGNIDPCLCTHLIYAFAGQNNNEITTIENWDTLYQAFN 81
QY 61 GLKKNKSOLKTLAIGGNWFGTAPFTAMYSTPNRQTFITSVIKELROYEPDGLDFDWEY 120
DB 82 GLKKNKTELKTLAIGGNWFGTAPFTAMYSTPNRQTFITSVIKELROYEPDGLDFDWEY 141
QY 121 PGSRGSPPQDKHLFTVLVQEMREAFQEAQKQINPKRLMVTAAVAGISNIQSGYIPLS 180
DB 142 PGSRGSPPKDKHLFSLVQEMKAFEEESTLNHI PRLLLTSTGAGFIDVIKSGYKIPLS 201
QY 181 QVLDIHWYTDLHGSWEGYTGENSEPLKYKPTDTSNAYLVNDYVNTWKONGAPAEKLI 240
DB 202 QSLDYIQVMTYDLHDPKDGTYGENSEPLKSPYDYGKADLNVDISITTYWKDHGAASEKLI 261
QY 241 VGFPTYGHNFIILSNPNTGIGAPTSGAGPAGPYAKESGIWAYVEICTFLKNGATQGDAP 300
DB 262 VGFPAHYHTFILDSPKGTIGAPTSTGPPGKYTDESGLLAYEVCITFLNEGTEWDAP 321
QY 301 QEVPIYQGNVWGYDNIKSFDIKAQWLKHNKFGGAMVWALDDFTGTFCNQKFPPLIS 360
DB 322 QEVPIYQGNVWGYDNIKSFDIKAQWLKHNKFGGAMVWALDDFTGTFCNQKFPPLIS 381
QY 361 TLKKAIGLOSASTAP 376
DB 382 TLKGLDNIHSASCKGP 397
```

DR ProDom; PD000471; Glyco_hydro_18; 1.
SQ SEQUENCE 402 AA; 44975 MW; D935389319A59905 CRC64;

Query Match
Best Local Similarity 55.2%; Score 1384; DB 11; Length 402;
Matches 250; Conservative 52; Mismatches 73; Indels 0; Gaps 0;

QY 1 YQLTCYFTNWAQYRGLGRFMPDNIDPCLCTHLIIYAFAGRONNEITTEIWNDDVTLYQAFN 60
DB 22 YQLMCYITSWAKDRPTEGSFKFNGNIDPCLCTHLIIYAFAGMKKNEITYLSEQDLRDYEALN 81

QY 61 GLKKNKSQLTKLLAIGGWNFGTAPTAMVSTPENRQTITSVKEFLROYEFGDLDFDWEY 120
DB 82 GLKDRNTELTLLAIGGWNFGTAPTAMVSTPENRQTITSVKEFLROYEFGDLDFDWEY 141

QY 121 PGRSGSPQDKHLFTLVQEMREAEQEAQKINPKRLMVTAAVAAGISNIOSGYEIPOLS 180
DB 142 PGRSGSPQDKHLFTLVQEMREAEQEAQKINPKRLMVTAAVAAGISNIOSGYEIPOLS 201

QY 181 QYLDYIHVMTYDLHGSWEGYTGNSPLYKYPTDTGNSNAYLVNVDYVMYWKONGAPAEKLI 240
DB 202 QSLDYIQVMTYDLHDPKNGYTGNSPLYKSPYDICKSADLVNDSIITYWKOHGAASEKLI 261

QY 241 VGFTYGHNFILSNPNTGIGAPTSGAGPAGPYAKESGIWAYEICTFLKNGATOGWDAP 300
DB 262 VGFPAYGHTFILDSPKNGIGDPTVSAGPPGKYTNEQGLLAYFEICTFLNEGATEIFDAT 321

QY 301 QEVPIYAGNVVGVYDNIKSPDIKAQMLKHNKFGAMWAIDLDDFTGTCNQGKFPPLIS 360
DB 322 QEVPIYAGNVVGVYDNIKSPDIKAQMLKHNKFGAMWAIDLDDFTGTCNQGKFPPLIS 381

QY 361 TLKKAALGLQSASCTA 375
DB 382 TLKRDNLNVHSASCKA 396

RESULT 13
Q8VH43 PRELIMINARY; PRT; 402 AA.

ID Q8VH43
AC Q8VH43
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DE Putative secretory protein.
GN YM2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C; TISSUE=STOMACH;
RA Yang S.C.; Su W.B.;
RT "Tissue-specific Expression of Ym2 Protein."
RL Submitted (DEC-2001) to the EMBL/GenBank/DDJ databases.
DR EMBL; AY065557; AAL57751.1; --
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
SQ SEQUENCE 402 AA; 44985 MW; D935382E89A5981A CRC64;

Query Match
Best Local Similarity 55.0%; Score 1379; DB 11; Length 402;
Matches 249; Conservative 52; Mismatches 74; Indels 0; Gaps 0;

QY 1 YQLTCYFTNWAQYRGLGRFMPDNIDPCLCTHLIIYAFAGRONNEITTEIWNDDVTLYQAFN 60
DB 22 YQLMCYITSWAKDRPTEGSFKFNGNIDPCLCTHLIIYAFAGMKKNEITYLSEQDLRDYEALN 81

QY 61 GLKKNKSQLTKLLAIGGWNFGTAPTAMVSTPENRQTITSVKEFLROYEFGDLDFDWEY 120
DB 82 GLKDRNTELTLLAIGGWNFGTAPTAMVSTPENRQTITSVKEFLROYEFGDLDFDWEY 141

QY 121 PGRSGSPQDKHLFTLVQEMREAEQEAQKINPKRLMVTAAVAAGISNIOSGYEIPOLS 180
DB 142 PGRSGSPQDKHLFTLVQEMREAEQEAQKINPKRLMVTAAVAAGISNIOSGYEIPOLS 201

QY 181 QYLDYIHVMTYDLHGSWEGYTGNSPLYKYPTDTGNSNAYLVNVDYVMYWKONGAPAEKLI 240
DB 202 QSLDYIQVMTYDLHDPKNGYTGNSPLYKSPYDICKSADLVNDSIITYWKOHGAASEKLI 261

QY 241 VGFTYGHNFILSNPNTGIGAPTSGAGPAGPYAKESGIWAYEICTFLKNGATOGWDAP 300
DB 262 VGFPAYGHTFILDSPKNGIGDPTVSAGPPGKYTNEQGLLAYFEICTFLNEGATEIFDAT 321

QY 301 QEVPIYAGNVVGVYDNIKSPDIKAQMLKHNKFGAMWAIDLDDFTGTCNQGKFPPLIS 360
DB 322 QEVPIYAGNVVGVYDNIKSPDIKAQMLKHNKFGAMWAIDLDDFTGTCNQGKFPPLIS 381

QY 361 TLKKAALGLQSASCTA 375
DB 382 TLKRDNLNVHSASCKA 396

RESULT 14
Q90W34 PRELIMINARY; PRT; 488 AA.

ID Q90W34
AC Q90W34
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Toad pancreatic chitinase (EC 3.2.1.14).
GN TPCASE.
OS Bufo japonicus (Japanese toad).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Bufonidae; Bufo.
OX NCBI_TaxID=8387;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PANCREAS;
RA Oshima H.;
RL Thesis (2001), Department of Department of Biology, Waseda University,
RL Tokyo, Japan.
DR EMBL; AJ345054; CAC87888.1; --
DR InterPro; IPR002086; Aldohyde_dehydr.
DR InterPro; IPR001579; Chitinase 18/2.
DR InterPro; IPR002557; Chitin bind PerA.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF01607; CBM_14; 1.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR PROSITE; PS00070; ALDEHYDE DEHYDR CYS; UNKNOWN_1.
DR PROSITE; PS01095; CHITINASE 18; UNKNOWN_1.
KW Glycosidase; Hydrolase; Signal.
FT SIGNAL 1 18
FT CHAIN 19 488 TOAD PANCREATIC CHITINASE.
SQ SEQUENCE 488 AA; 54319 MW; E75E1AD3CB2B4919 CRC64;

Query Match
Best Local Similarity 53.5%; Score 1340.5; DB 13; Length 488;
Matches 249; Conservative 58; Mismatches 133; Indels 43; Gaps 4;

QY 1 YQLTCYFTNWAQYRGLGRFMPDNIDPCLCTHLIIYAFAGRONNEITTEIWNDDVTLYQAFN 60
DB 19 YKLVCFYFTNWSQYRPGDQKGVPGNIDPCLCTHLVYAFATMNEHKIAPYEWNDVLYKQFN 78

QY 61 GLKKNKSQLTKLLAIGGWNFGTAPTAMVSTPENRQTITSVKEFLROYEFGDLDFDWEY 120
DB 79 DLKQKNKLVTLAIGGWNFGTAPTAMVSTPENRQTITSVKEFLROYEFGDLDFDWEY 138

QY 121 PGRSGSPQDKHLFTLVQEMREAEQEAQKINPKRLMVTAAVAAGISNIOSGYEIPOLS 180
DB 139 PGRSGSPQDKHLFTLVQEMREAEQEAQKINPKRLMVTAAVAAGISNIOSGYEIPOLS 198

QY 181 QYLDYIHVMTYDLHGSWEGYTGNSPLYKYPTDTGNSNAYLVNVDYVMYWKONGAPAEKLI 240
DB 198 QYLDYIHVMTYDLHGSWEGYTGNSPLYKYPTDTGNSNAYLVNVDYVMYWKONGAPAEKLI 240

Db 199 QLLDFISVMTYDFHGWDTQSGHNSPLCKSGSTDYGLQYNIHFAMVYWKNGAPAEKLL 258
QY 241 VGPPTGHNFIILNPNTGIGAPTSGAGPAGPYAKESGIWAYYEICTFLKNGATQGWDA 300
Db 259 LGPPTVGRTRNPNPCNDVGIPVSGAGSAGPYTREAGFWAYYEICTWL-SGSTVKWIPD 317
QY 301 QEVPYAYQGVVYGYDNIKSFDIKAQWLKHNKFGGAMWAIDLDDFTGTCNQGKPEPLIS 360
Db 318 QRPVYACKSNEWGFDNQESYECKVFLKESGFGGAMWAIDLDDFEGRCNQGRYPLIN 377
QY 361 TLKAL-----GLQ-----SASCTAPAOPIEPIITAAPSSGG 391
Db 378 HLKSLLEGSTVNCBICGSIPTSPSATTITTTTAKPDCCTTPBPVTPPVPPV--- 434
QY 392 NGSAGSSSSGSGSGCFCAVRANGLYPVANNRNFHVCVNGVYQQNCOAGLVFDTSDC 451
Db 435 -----IDVDPNFCVEKTDGLHVNPLNTKFKYICANGRTYSMKCADGLVFOASCNC 484
QY 452 CNW 454
Db 485 CNW 487

RESULT 15

Q13231
ID Q13231 PRELIMINARY; PRT; 466 AA.
AC Q13231;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Chitotriosidase precursor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96064695; PubMed=7592832;
RA Boot R.G., Renkema G.H., Strijland A., van Zonneveld A.J., Aerts J.M.;
RT "Cloning of a cDNA encoding chitotriosidase, a human chitinase
produced by macrophages.";
RL J. Biol. Chem. 270:26252-26256(1995).
CC -!- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
HYDROLASES).
CC EMBL: U29615; AAC50246.1; -;
DR InterPro: IPR001579; Chitinase 18/2.
DR InterPro: IPR002557; Chitin_bind_Pera.
DR InterPro: IPR001223; Glyco_Hydro_18.
DR Pfam: PF01607; CBM 14; 1.
DR ProDom: PD00704; Glyco_hydro_18; 1.
DR SMART: SM00494; ChnB2; 1.
DR PROSITE: PS01095; CHITINASE_18; 1.
KW Glycosidase; Hydrolase; Signal.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 466 CHITOTRIOSIDASE.
SQ SEQUENCE 466 AA; 51681 MW; B4312D1E885E386D CRC64;

Query Match 52.8%; Score 1323; DB 4; Length 466;
Best Local Similarity 52.2%; Pred. No. 8.7e-87;
Matches 240; Conservative 70; Mismatches 126; Indels 24; Gaps 3;

QY 2 QLTCTFTNNAQYRPGIGREFMPONIDPCLCTHLIYAPAGRONNEITTEWNDVTLQAFNG 61
Db 23 KLVCTFTNNAQYRQGEARFLPKDPLSLCTHLIYAPAGTNHQLSTEWNDVTLQAFNG 82
QY 62 LKKNLSQLKTLAIGGNFTGAPFTAMVSTPENRQFTTSVIKFLRQYFDFGLDFDWEYP 121
Db 83 LKKNPFLKTLAIGGNFTGAPFTAMVSTPENRQFTTSVIKFLRQYFDFGLDFDWEYP 142
QY 122 GSGSPQDKHLFTVLVQEMREAFQEAQKINPRLMVTAAVAAGISNIQSVEYFPLSQ 181
Db 143 GSGSPQDKERFTTLVQDLANAFQEAQTSKGERLLLSAAVPAGQTVYDAGYVDKIAQ 202

QY 182 YLDYIHVMTYDLHGSWEGVTGNSPLKYPTDTGSNAYLVNDYVNMVYWKNGAPAEKLTIV 241
Db 203 NUDFVNLWAYDFHGSWEKVTGNSPLKYRQESGAAASLVNDAAYQOHLQKTPASKLIL 262
QY 242 GPPTYGHNFIILNPNTGIGAPTSGAGPAGPYAKESGIWAYYEICTFLKNGATQGWDA 301
Db 263 GWPTYGRSFTLASSSDTRVGPATSGTGPPTKGGMLAYYEVCSW--KGATKQRIQDQ 320
QY 302 EYPYAYQGVVYGYDNIKSFDIKAQWLKHNKFGGAMWAIDLDDFTGTCNQGKPEPLIS 361
Db 321 KVPYIFRDNQWGFODVESFKTKVSYLKQKGLGGAMVWALDLDLDFAGFSCNQGRYPLIQ 380
QY 362 LKAL-----GLQASCTAPAOPIEPIITAAPSGNGSGSSSGSGSGGFCVAVRAN 414
Db 381 LRQELSLPYLPSGTPELEVPKGPQSEP-----EKGPSFGQDTFCQKAD 425
QY 415 GLYPVANNRNFHVCVNGVYQQNCOAGLVFDTSDCNCW 454
Db 426 GLYPNPRSSSFYSCAAGRLFOQSCPTGLVFSNSCKCCTW 465

Search completed: June 29, 2003, 21:01:48
Job time : 26.5769 secs

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	1323	52.8	466	2	US-08-486-839-4	Sequence 4, Appli
2	1323	52.8	466	3	US-09-151-011-4	Sequence 4, Appli
3	1323	52.8	466	4	US-09-039-198A-2	Sequence 2, Appli
4	1323	52.8	466	4	US-09-343-623-4	Sequence 4, Appli
5	1323	52.8	466	4	US-08-877-599-2	Sequence 2, Appli
6	1323	52.8	466	4	US-08-267-574-2	Sequence 2, Appli
7	1317	52.6	466	4	US-09-039-198A-4	Sequence 4, Appli
8	1317	52.6	466	4	US-08-877-599-4	Sequence 4, Appli
9	1317	52.6	466	4	US-09-267-574-4	Sequence 4, Appli
10	1177	47.0	373	4	US-09-039-198A-14	Sequence 14, Appli
11	1177	47.0	373	4	US-09-039-198A-15	Sequence 15, Appl
12	1177	47.0	373	4	US-08-877-599-14	Sequence 14, Appl
13	1177	47.0	373	4	US-08-877-599-15	Sequence 15, Appl
14	1177	47.0	373	4	US-09-267-574-14	Sequence 14, Appl
15	1177	47.0	373	4	US-09-267-574-15	Sequence 15, Appl
16	1173	46.8	387	2	US-08-486-839-6	Sequence 6, Appli
17	1173	46.8	387	3	US-09-151-011-6	Sequence 6, Appli
18	1173	46.8	387	4	US-09-343-623-6	Sequence 6, Appli
19	970.5	38.7	385	2	US-08-694-915-2	Sequence 2, Appli
20	970.5	38.7	416	2	US-08-694-915-4	Sequence 4, Appli
21	968.5	38.6	383	4	US-09-459-749D-17	Sequence 17, Appli
22	863	34.4	554	2	US-08-524-051-2	Sequence 2, Appli
23	863	34.4	554	3	US-09-052-778-16	Sequence 16, Appl
24	798.5	31.9	559	4	US-09-545-814-14	Sequence 14, Appl
25	798.5	31.9	583	4	US-09-545-814-2	Sequence 2, Appli
26	798.5	31.9	583	4	US-09-545-814-5	Sequence 5, Appli
27	794.5	31.7	635	4	US-09-545-814-32	Sequence 32, Appl

Query Match 52.8%; Score 1323; DB 3; Length 466;
Best Local Similarity 52.2%; Pred. No. 1.4e-111;
Matches 240; Conservative 70; Mismatches 126; Indels 24; Gaps 3;
QY 62 LKKNKSLKTLAIGWNGFTAPPTAMVSTPENRQTFITSVIKLROYEFDGLDFDWEYP 121
DB 83 LKKNKSLKTLAIGWNGFTAPPTAMVSTPENRQTFITSVIKLROYEFDGLDFDWEYP 142
QY 122 GSRGSPQDKHLFTVLVQEMREAFQEAQKQINPKRLMVTAAVAAGISNIQSIEIPOLSQ 181
DB 143 GSGSPAVDKERFTTLVQDLANAFQEAQTSKGERLLLSAAVPAQOTYVDAGYEVDKIAQ 202
QY 182 YLDYIHVMTYDLHGSWEGVTGENSPLYKPYPTDTGNSNAYLVNVDYVMYKDNKNGAPAEKLI 241
DB 203 NLDYFVNLMAIDFHGSWEKVTGHNPLKQKQESGAASLVNDAVQOVLKQKGPASKLIL 262
QY 242 GFTYGHNFILSNPNTGIGAPTSGAGPAGPYAKESGIWAYVEICTFLKNGATOGWDAPO 301
DB 263 GMFTYGRSFTLASSSDTRVGAPATGSGTGPFTKEGMLAYEVCWS--KGATKQRIQOQ 320
QY 302 EYVAYQGNVWVGYDNIKSFIDIKAKQMLKHNKFGAMVWALDLDFTGTCNQGKFLIST 361
DB 321 KVPYIFRDQWVGFDVESFRTKVSYLKQKGLGAMVWALDLDFTGTCNQGKFLIST 380
QY 362 LKAL-----GLOSASCTAPAOPIEPIAAPSNGSGSGSSGSGSGGSCGFCAVRAN 414
DB 381 LRQELSLPLPSGTPELEVPKQPSSEP-----EHGSPSQDFTFCQKAD 425
QY 415 GLYPVANNRNFHVCNGVTVQONCOAGLVFDTSCDCCNW 454
DB 426 GLYPNRRSSFFSCAAGRLFQOSCTGLVFSNCKCCTW 465

RESULT 2
US-09-151-011-4
; Sequence 4, Application US/09151011
; Patent No. 6057142
; GENERAL INFORMATION:
; APPLICANT: Hoffmann & Baron, LLP
; STREET: 6900 Jericho Turnpike
; CITY: Syosset
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11791
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/151,011
; FILING DATE: 10 - September - 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Morris, Robert C.
; REGISTRATION NUMBER: 42,910
; REFERENCE/DOCKET NUMBER: 294-32 DIV
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 466 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
US-09-151-011-4

Query Match 52.8%; Score 1323; DB 3; Length 466;
Best Local Similarity 52.2%; Pred. No. 1.4e-111;
Matches 240; Conservative 70; Mismatches 126; Indels 24; Gaps 3;
QY 2 QLTCTFTWQVPRGLGRPMNDIPCLCTHLIYAFAGRONNEITTIWVTLVQAFNG 61
DB 23 KLVCFYFTWQVPRGLGRPMNDIPCLCTHLIYAFAGRONNEITTIWVTLVQAFNG 82
QY 62 LKKNKSLKTLAIGWNGFTAPPTAMVSTPENRQTFITSVIKLROYEFDGLDFDWEYP 121
DB 83 LKKNKSLKTLAIGWNGFTAPPTAMVSTPENRQTFITSVIKLROYEFDGLDFDWEYP 142
QY 122 GSRGSPQDKHLFTVLVQEMREAFQEAQKQINPKRLMVTAAVAAGISNIQSIEIPOLSQ 181
DB 143 GSGSPAVDKERFTTLVQDLANAFQEAQTSKGERLLLSAAVPAQOTYVDAGYEVDKIAQ 202
QY 182 YLDYIHVMTYDLHGSWEGVTGENSPLYKPYPTDTGNSNAYLVNVDYVMYKDNKNGAPAEKLI 241
DB 203 NLDYFVNLMAIDFHGSWEKVTGHNPLKQKQESGAASLVNDAVQOVLKQKGPASKLIL 262
QY 242 GFTYGHNFILSNPNTGIGAPTSGAGPAGPYAKESGIWAYVEICTFLKNGATOGWDAPO 301
DB 263 GMFTYGRSFTLASSSDTRVGAPATGSGTGPFTKEGMLAYEVCWS--KGATKQRIQOQ 320
QY 302 EYVAYQGNVWVGYDNIKSFIDIKAKQMLKHNKFGAMVWALDLDFTGTCNQGKFLIST 361
DB 321 KVPYIFRDQWVGFDVESFRTKVSYLKQKGLGAMVWALDLDFTGTCNQGKFLIST 380
QY 362 LKAL-----GLOSASCTAPAOPIEPIAAPSNGSGSGSSGSGSGGSCGFCAVRAN 414
DB 381 LRQELSLPLPSGTPELEVPKQPSSEP-----EHGSPSQDFTFCQKAD 425
QY 415 GLYPVANNRNFHVCNGVTVQONCOAGLVFDTSCDCCNW 454
DB 426 GLYPNRRSSFFSCAAGRLFQOSCTGLVFSNCKCCTW 465

RESULT 3
US-09-039-198A-2
; Sequence 2, Application US/09039198A
; Patent No. 6200951
; GENERAL INFORMATION:
; APPLICANT: Gray, Patrick W.
; APPLICANT: Tjoelker, Larry W.
; TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive/6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/039,198A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 27866/34391
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 466 amino acids
; TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-039-198A-2

Query Match 52.8%; Score 1323; DB 4; Length 466;
Best Local Similarity 52.2%; Pred. No. 1.4e-111;
Matches 240; Conservative 70; Mismatches 126; Indels 24; Gaps 3;

QY 2 QUTCYFTNMAQYRPGGLGRFMPDNIDPCLCTHLIYAFAGQNNNEITTEIWNQVTLQAFNG 61
DB 23 KLVCFYFTNMAQYRQGEARFLPKDLPSLCTHLIYAFAGMTNHLSTTEWDETLYQEFNG 82
QY 62 LKVNQSOLKTLAIGGNWFGTAPFTAMVSTPENROTFTITSVIFKFLROYEFDGLDDEWYEP 121
DB 83 LKKNMPKLTLLAIGGNWFGTQKFTDMVATANNROTFTVNSAIRFLKYSFDFGLDDEWYEP 142
QY 122 GSGSPQDKHLFTVLVOEMREAFQEAQKQINPKRLMVTAAVAAGISNIQSGYEIPOLSQ 181
DB 143 GSGSPAVDKERFTTLVQDLANAFQEAQTSKGERLLLSAAVPAQOTYYDAGYEVDKIAQ 202
QY 182 YLDYIHVMTYDLHGSWEGYTGNSPLYKYPTDTSNAYLNVDYVMYKNDGAPAEKLIIV 241
DB 203 NLDVFNLMAYDFHGSWEKVTGHNSPLYKQESGAAASLNVDAAVQOMLQKGTAPASKLIL 262
QY 242 GPPTYGHNFILSNPTGIGAPTSAGPAGPYAKESGIWAYEICTFLKNGATQGWDAPO 301
DB 263 GMPTYGSRFTLASSSDTRVGA PATSGTGPPTKEGMLAYEVCWS--KGATKQRIQDQ 320
QY 302 EYPYAYQGNVWYGYDNIKSFDIKAQWLKHNKFGGAMVWALDLDFTGTFCNOGKPELIST 361
DB 321 KVPYIFRDNQWVGDFDVESEFKTKVSLKQGLGGAMVWALDLDFAFGSCNQRYPLOT 380
QY 362 LKAL-----GLOSASCTAPAEPTPIAAPSNGSGSGSSGSGSGGFCACAVRAN 414
DB 381 LKQELSLPLPSGTPELEVPKQPSSEP-----EHGPSQDQTFQCGKAD 425
QY 415 GLYPVANNRNPFWHCUNGVTYQONCOAGLVFDTSCDCCNW 454
DB 426 GLYPNPRSSFFSACAAGRLFQOCSCTGLVFSNSCKCCTW 465

RESULT 4

US-09-343-623-4
Sequence 4, Application US/09343623
Patent No. 6303118
GENERAL INFORMATION:

APPLICANT: A human chitinase, its recombinant
TITLE OF INVENTION: production, its use for decomposing chitin, its use
in therapy or prophylaxis against infection diseases.

NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:

ADDRESSEE: Hoffmann & Baron
CITY: Jericho
STATE: New York
COUNTRY: United States of America
ZIP: 11758

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/343,623
FILING DATE:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/08/486,839
FILING DATE: 07-June-1995
ATTORNEY/AGENT INFORMATION:
NAME: Baron, Ronald J.
REGISTRATION NUMBER: 29,281

REFERENCE/DOCKET NUMBER: 294-26
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 822-3550
TELEFAX: (516) 822-3582

INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 466 amino acids

TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-09-343-623-4

Query Match 52.8%; Score 1323; DB 4; Length 466;
Best Local Similarity 52.2%; Pred. No. 1.4e-111;
Matches 240; Conservative 70; Mismatches 126; Indels 24; Gaps 3;

QY 2 QUTCYFTNMAQYRPGGLGRFMPDNIDPCLCTHLIYAFAGQNNNEITTEIWNQVTLQAFNG 61
DB 23 KLVCFYFTNMAQYRQGEARFLPKDLPSLCTHLIYAFAGMTNHLSTTEWDETLYQEFNG 82
QY 62 LKVNQSOLKTLAIGGNWFGTAPFTAMVSTPENROTFTITSVIFKFLROYEFDGLDDEWYEP 121
DB 83 LKKNMPKLTLLAIGGNWFGTQKFTDMVATANNROTFTVNSAIRFLKYSFDFGLDDEWYEP 142
QY 122 GSGSPQDKHLFTVLVOEMREAFQEAQKQINPKRLMVTAAVAAGISNIQSGYEIPOLSQ 181
DB 143 GSGSPAVDKERFTTLVQDLANAFQEAQTSKGERLLLSAAVPAQOTYYDAGYEVDKIAQ 202
QY 182 YLDYIHVMTYDLHGSWEGYTGNSPLYKYPTDTSNAYLNVDYVMYKNDGAPAEKLIIV 241
DB 203 NLDVFNLMAYDFHGSWEKVTGHNSPLYKQESGAAASLNVDAAVQOMLQKGTAPASKLIL 262
QY 242 GPPTYGHNFILSNPTGIGAPTSAGPAGPYAKESGIWAYEICTFLKNGATQGWDAPO 301
DB 263 GMPTYGSRFTLASSSDTRVGA PATSGTGPPTKEGMLAYEVCWS--KGATKQRIQDQ 320
QY 302 EYPYAYQGNVWYGYDNIKSFDIKAQWLKHNKFGGAMVWALDLDFTGTFCNOGKPELIST 361
DB 321 KVPYIFRDNQWVGDFDVESEFKTKVSLKQGLGGAMVWALDLDFAFGSCNQRYPLOT 380
QY 362 LKAL-----GLOSASCTAPAEPTPIAAPSNGSGSGSGSGSGGFCACAVRAN 414
DB 381 LKQELSLPLPSGTPELEVPKQPSSEP-----EHGPSQDQTFQCGKAD 425
QY 415 GLYPVANNRNPFWHCUNGVTYQONCOAGLVFDTSCDCCNW 454
DB 426 GLYPNPRSSFFSACAAGRLFQOCSCTGLVFSNSCKCCTW 465

RESULT 5

US-08-877-599-2
Sequence 2, Application US/08877599
Patent No. 6372212
GENERAL INFORMATION:

APPLICANT: Gray, Patrick W.
TITLE OF INVENTION: Chitinase Materials and Methods

NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/877,599

```

RESULT 7
US-09-039-198A-4
; Sequence 4, Application US/09039198A
; Patent No. 6200951
; GENERAL INFORMATION:
; APPLICANT: Gray, Patrick W.
; APPLICANT: Tjoelker, Larry W.
; TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Bor
; STREET: 233 South Wacker Drive/6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/039.198A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 27866/34391
; TELECOMMUNICATION INFORMATION:

```

GENERAL INFORMATION:
 APPLICANT: Gray, Patrick W.
 TITLE OF INVENTION: Chitnase Materials and Methods
 NUMBER OF SEQUENCES: 17
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
 STREET: 6300 Sears Tower, 233 South Wacker Drive
 CITY: Chicago
 STATE: Illinois
 COUNTRY: United States of America
 ZIP: 60606-6402
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/877,599
 FILING DATE:
 CLASSIFICATION: 514
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/663,618

US-09-267-574-4

Query Match 52.6%; Score 1317; DB 4; Length 466;
Best Local Similarity 52.0%; Pred. No. 4.8e-111;
Matches 239; Conservative 70; Mismatches 127; Indels 24; Gaps 3;

QY 2 QLTCTFTNNAQYRPGGLGRFMPDNIDPCLCTHLIYAFAGRONNEITIEWNDVLYQAFNG 61
DB 23 KLVCYFTNNAQYRQGEARFLPKDLPSLCTHLIYAFAGTNHQLSTTEWNETLYQEFNG 82
QY 62 LKKNKSQLKTLAIGWNFGTAPFTAMVSTPENROTFTITSVTKFLROYEFDGLDFDWEYP 121
DB 83 LKKNPKLTLAIGWNFGTAPFTAMVSTPENROTFTITSVTKFLROYEFDGLDFDWEYP 142
QY 122 GSRGSPQDKHLFTVLVQEMREAFQEAQINPKRLMVTAAVAAGISNIOGVEIPOLSO 181
DB 143 GSQSPAVDKERFTTLVQDLANAFQEAQTSKGERULLLSAAVPAQOTYVDAGYEVDKIAQ 202
QY 182 YLDYIHVMTYDLHGSWEGTGENSPLYKYPTDTGNSNAYLVNDYVMNWKDNGAPAEKLI 241
DB 203 NLDVFNLMAYDFHGSWEKVTGHNSPLYKQESGAAASLVNDAVQOVLQKGTTPASKLIL 262
QY 242 GPTYGHNFILSPNTGTGAPTSGAGPAGPYAKESGIWAYEICTFLKNGATQGWDAPO 301
DB 263 GMPTGRSFTLASSSDTRVGAPATGSGTTPGFTKEGMLAYEVCWS--KGATKQRIQDQ 320
QY 302 EYPYAYQGNVWVGYDNIKSFIDIAKQWLKHNKFGAMWALDLDFTGTFCNQKPELIST 361
DB 321 KVPYIFRDQWGVDFDVESEFKTKVSYLKQKGLGGAMWALDLDFTGTFCNQKPELIST 380
QY 362 LKALGL 368
DB 360 LRQELSL 366

RESULT 10

US-09-039-198A-14
Sequence 14, Application US/09039198A
Patent No. 6200951
GENERAL INFORMATION:
APPLICANT: Gray, Patrick W.
APPLICANT: Tjoelker, Larry W.
TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive/6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/039,198A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Rin-Laures, Li-Hsien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 27866/34391
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:

LENGTH: 373 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-039-198A-14

Query Match 47.0%; Score 1177; DB 4; Length 373;
Best Local Similarity 56.9%; Pred. No. 1.7e-98;
Matches 209; Conservative 63; Mismatches 93; Indels 2; Gaps 1;

QY 2 QLTCTFTNNAQYRPGGLGRFMPDNIDPCLCTHLIYAFAGRONNEITIEWNDVLYQAFNG 61
DB 2 KLVCYFTNNAQYRQGEARFLPKDLPSLCTHLIYAFAGTNHQLSTTEWNETLYQEFNG 61
QY 62 LKKNKSQLKTLAIGWNFGTAPFTAMVSTPENROTFTITSVTKFLROYEFDGLDFDWEYP 121
DB 62 LKKNPKLTLAIGWNFGTAPFTAMVSTPENROTFTITSVTKFLROYEFDGLDFDWEYP 121
QY 122 GSRGSPQDKHLFTVLVQEMREAFQEAQINPKRLMVTAAVAAGISNIOGVEIPOLSO 181
DB 122 GSQSPAVDKERFTTLVQDLANAFQEAQTSKGERULLLSAAVPAQOTYVDAGYEVDKIAQ 181
QY 182 YLDYIHVMTYDLHGSWEGTGENSPLYKYPTDTGNSNAYLVNDYVMNWKDNGAPAEKLI 241
DB 182 NLDVFNLMAYDFHGSWEKVTGHNSPLYKQESGAAASLVNDAVQOVLQKGTTPASKLIL 241
QY 242 GPTYGHNFILSPNTGTGAPTSGAGPAGPYAKESGIWAYEICTFLKNGATQGWDAPO 301
DB 242 GMPTGRSFTLASSSDTRVGAPATGSGTTPGFTKEGMLAYEVCWS--KGATKQRIQDQ 299
QY 302 EYPYAYQGNVWVGYDNIKSFIDIAKQWLKHNKFGAMWALDLDFTGTFCNQKPELIST 361
DB 300 KVPYIFRDQWGVDFDVESEFKTKVSYLKQKGLGGAMWALDLDFTGTFCNQKPELIST 359
QY 362 LKALGL 368
DB 360 LRQELSL 366

RESULT 11

US-09-039-198A-15
Sequence 15, Application US/09039198A
Patent No. 6200951
GENERAL INFORMATION:
APPLICANT: Gray, Patrick W.
APPLICANT: Tjoelker, Larry W.
TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive/6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/039,198A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Rin-Laures, Li-Hsien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 27866/34391
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:
LENGTH: 373 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-039-198A-15

Query Match 47.0%; Score 1177; DB 4; Length 373;
Best Local Similarity 56.9%; Pred. No. 1.7e-98;
Matches 209; Conservative 63; Mismatches 93; Indels 2; Gaps 1;

QY 2 QLTCTFTNAQYRPGLGFRMPDNIDPCLCTHLIYAFAGRONNEITTIWNDVTLYOAFNG 61
DB 2 KLVCVFTNAQYRQGEARFLPKDLPSCSLCTHLIYAFAGMTNHQSLTTEWNDVTLYOAFNG 61
QY 62 LKKNKSQLKTLTLLAIGWNFGTAPFTAMVSTPENROTFTTSVTKFLROYEFDGLDDEWYEP 121
DB 62 LKKNKSQLKTLTLLAIGWNFGTAPFTAMVSTPENROTFTTSVTKFLROYEFDGLDDEWYEP 121
QY 122 GSRGSPPODKHLFTVLVQEMREAFQEAQKINPKRLMTVAAGISNIOGVEIPOLSQ 181
DB 122 GSRGSPPODKHLFTVLVQEMREAFQEAQKINPKRLMTVAAGISNIOGVEIPOLSQ 181
QY 182 YLDYIHVMTYDLHGSWEVGTGNSPLYKYPTDTGNSNAYLNVDYVWYKNDGAPAEKLIV 241
DB 182 YLDYIHVMTYDLHGSWEVGTGNSPLYKYPTDTGNSNAYLNVDYVWYKNDGAPAEKLIV 241
QY 242 GPPTGHNFIILNSPNTGIGAPTSGAGPAGPVAKESGIWAYVEICTFLKNGATQGWDAPO 301
DB 242 GPPTGHNFIILNSPNTGIGAPTSGAGPAGPVAKESGIWAYVEICTFLKNGATQGWDAPO 301
QY 302 EYPYVQGVNVVGYDNIKSFDIKAQWLKHNKFGGAMVWALDLDFTGTFCNQGKPELIST 361
DB 302 EYPYVQGVNVVGYDNIKSFDIKAQWLKHNKFGGAMVWALDLDFTGTFCNQGKPELIST 361
QY 362 LKKAALGL 368
DB 362 LKKAALGL 368

RESULT 12

US-08-877-599-14
Sequence 14, Application US/08877599
Patent No. 6372212

GENERAL INFORMATION:

APPLICANT: Gray, Patrick W.
TITLE OF INVENTION: Chitinase Materials and Methods
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSES: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/877,599
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/663,618
FILING DATE: 14-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Rin-Laures, Li-Hsien
REGISTRATION NUMBER: 33,547
REFERENCES/DOCKET NUMBER: 27866/33994
TELECOMMUNICATION INFORMATION:

TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 373 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-877-599-14

Query Match 47.0%; Score 1177; DB 4; Length 373;
Best Local Similarity 56.9%; Pred. No. 1.7e-98;
Matches 209; Conservative 63; Mismatches 93; Indels 2; Gaps 1;

QY 2 QLTCTFTNAQYRPGLGFRMPDNIDPCLCTHLIYAFAGRONNEITTIWNDVTLYOAFNG 61
DB 2 KLVCVFTNAQYRQGEARFLPKDLPSCSLCTHLIYAFAGMTNHQSLTTEWNDVTLYOAFNG 61
QY 62 LKKNKSQLKTLTLLAIGWNFGTAPFTAMVSTPENROTFTTSVTKFLROYEFDGLDDEWYEP 121
DB 62 LKKNKSQLKTLTLLAIGWNFGTAPFTAMVSTPENROTFTTSVTKFLROYEFDGLDDEWYEP 121
QY 122 GSRGSPPODKHLFTVLVQEMREAFQEAQKINPKRLMTVAAGISNIOGVEIPOLSQ 181
DB 122 GSRGSPPODKHLFTVLVQEMREAFQEAQKINPKRLMTVAAGISNIOGVEIPOLSQ 181
QY 182 YLDYIHVMTYDLHGSWEVGTGNSPLYKYPTDTGNSNAYLNVDYVWYKNDGAPAEKLIV 241
DB 182 YLDYIHVMTYDLHGSWEVGTGNSPLYKYPTDTGNSNAYLNVDYVWYKNDGAPAEKLIV 241
QY 242 GPPTGHNFIILNSPNTGIGAPTSGAGPAGPVAKESGIWAYVEICTFLKNGATQGWDAPO 301
DB 242 GPPTGHNFIILNSPNTGIGAPTSGAGPAGPVAKESGIWAYVEICTFLKNGATQGWDAPO 301
QY 302 EYPYVQGVNVVGYDNIKSFDIKAQWLKHNKFGGAMVWALDLDFTGTFCNQGKPELIST 361
DB 302 EYPYVQGVNVVGYDNIKSFDIKAQWLKHNKFGGAMVWALDLDFTGTFCNQGKPELIST 361
QY 362 LKKAALGL 368
DB 362 LKKAALGL 368

RESULT 13

US-08-877-599-15
Sequence 15, Application US/08877599
Patent No. 6372212

GENERAL INFORMATION:

APPLICANT: Gray, Patrick W.
TITLE OF INVENTION: Chitinase Materials and Methods
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSES: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/877,599
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/663,618
FILING DATE: 14-JUN-1996
ATTORNEY/AGENT INFORMATION:

NAME: Rin-Laures, Li-Hsien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 27866/33994
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:
LENGTH: 373 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide

US-08-877-599-15

Query Match 47.0%; Score 1177; DB 4; Length 373;
Best Local Similarity 56.9%; Pred. No. 1.7e-98;
Matches 209; Conservative 63; Mismatches 93; Indels 2; Gaps 1;

QY 2 QLTCTNNAQYRPGGLRFPMDNIDPCLCCTHLIYAFAGRONNEITTIENNDVTLQAENG 61
DB 2 KLVCTFTNNAQYRQGEARFLPKDLPSLCTHLIYAFAGMTNHLSTTEWNETLYQEFNG 61
QY 62 LKNKNSQLKTLAIGGNWFTAPFTAMVSTPENRQFTTTSVIKFLRQYFDFGLDFDWEYP 121
DB 62 LKKNMPLKTLAIGGNWFTQKFTDMVATANNRQTFVNSAIRFLRKYSFDFGLDLDWEYP 121
QY 122 GSRGSPPODKHLFTVLVOEMREAFQEAQKINPKRLMVTAAVAAGISNIQSGEYEPQLSQ 181
DB 122 GSQGSFAVDKERFTTLVQDLANAFQGEAQTSGKERLLLSAAVPAQGTVDAGYEVDKIAQ 181
QY 182 YLDYIHVMYDYLHGSWEGYTGENSEPLYKYPTDGTGNSAYLNVDYVMYKNDGAPAEKLI 241
DB 182 NLDVFNLMAYDFHGSWEKVTGHNSPLYKQESGAAASLNVDAAVQOMLQKGTGPASKLIL 241
QY 242 GFTYGHNFILSNPNTGIGAPTSAGPAGPYAKESGIWAYEICTFLKNGATQGWDAPO 301
DB 242 GMPTYGRSFTLASSSDTRVGAPATGSGTGPFTKEGMLAYEVCWS--KGATKQRIQDQ 299
QY 302 EYPYAVQGNVWGYDNIKSFEDIKAOMLKHNKFGGAMVWALDLDFTGTCNQGKPLIST 361
DB 300 KVPYIFRDNQWGFDDVESFTKVSILKQGLGGAMVWALDLDPAFSCNQGRIPLIQT 359
QY 362 LKALGL 368
DB 360 LRQELSL 366

RESULT 14

US-09-267-574-14

Sequence 14, Application US/09267574

Patent No. 639571

GENERAL INFORMATION:

APPLICANT: Gray, Patrick W.

APPLICANT: Tjoelker, Larry W.

TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS

FILE REFERENCE: 27866/35407

CURRENT APPLICATION NUMBER: US/09/267,574

CURRENT FILING DATE: 1999-03-12

EARLIER APPLICATION NUMBER: 09/039,198

EARLIER FILING DATE: 1998-03-12

NUMBER OF SEQ ID NOS: 39

SOFTWARE: Patent in Ver. 2.0

SEQ ID NO 14

LENGTH: 373

TYPE: PRT

ORGANISM: Homo sapiens

US-09-267-574-14

Query Match

Best Local Similarity 56.9%; Pred. No. 1.7e-98;
Matches 209; Conservative 63; Mismatches 93; Indels 2; Gaps 1;

QY 2 QLTCTNNAQYRPGGLRFPMDNIDPCLCCTHLIYAFAGRONNEITTIENNDVTLQAENG 61
DB 2 KLVCTFTNNAQYRQGEARFLPKDLPSLCTHLIYAFAGMTNHLSTTEWNETLYQEFNG 61
QY 62 LKNKNSQLKTLAIGGNWFTAPFTAMVSTPENRQFTTTSVIKFLRQYFDFGLDFDWEYP 121
DB 62 LKKNMPLKTLAIGGNWFTQKFTDMVATANNRQTFVNSAIRFLRKYSFDFGLDLDWEYP 121
QY 122 GSRGSPPODKHLFTVLVOEMREAFQEAQKINPKRLMVTAAVAAGISNIQSGEYEPQLSQ 181
DB 122 GSQGSFAVDKERFTTLVQDLANAFQGEAQTSGKERLLLSAAVPAQGTVDAGYEVDKIAQ 181
QY 182 YLDYIHVMYDYLHGSWEGYTGENSEPLYKYPTDGTGNSAYLNVDYVMYKNDGAPAEKLI 241
DB 182 NLDVFNLMAYDFHGSWEKVTGHNSPLYKQESGAAASLNVDAAVQOMLQKGTGPASKLIL 241
QY 242 GFTYGHNFILSNPNTGIGAPTSAGPAGPYAKESGIWAYEICTFLKNGATQGWDAPO 301
DB 242 GMPTYGRSFTLASSSDTRVGAPATGSGTGPFTKEGMLAYEVCWS--KGATKQRIQDQ 299
QY 302 EYPYAVQGNVWGYDNIKSFEDIKAOMLKHNKFGGAMVWALDLDFTGTCNQGKPLIST 361
DB 300 KVPYIFRDNQWGFDDVESFTKVSILKQGLGGAMVWALDLDPAFSCNQGRIPLIQT 359
QY 362 LKALGL 368
DB 360 LRQELSL 366

RESULT 15

US-09-267-574-15

Sequence 15, Application US/09267574

Patent No. 639571

GENERAL INFORMATION:

APPLICANT: Gray, Patrick W.

APPLICANT: Tjoelker, Larry W.

TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS

FILE REFERENCE: 27866/35407

CURRENT APPLICATION NUMBER: US/09/267,574

CURRENT FILING DATE: 1999-03-12

EARLIER APPLICATION NUMBER: 09/039,198

EARLIER FILING DATE: 1998-03-12

NUMBER OF SEQ ID NOS: 39

SOFTWARE: Patent in Ver. 2.0

SEQ ID NO 15

LENGTH: 373

TYPE: PRT

ORGANISM: Homo sapiens

US-09-267-574-15

Query Match 47.0%; Score 1177; DB 4; Length 373;

Best Local Similarity 56.9%; Pred. No. 1.7e-98;

Matches 209; Conservative 63; Mismatches 93; Indels 2; Gaps 1;

QY 2 QLTCTNNAQYRPGGLRFPMDNIDPCLCCTHLIYAFAGRONNEITTIENNDVTLQAENG 61
DB 2 KLVCTFTNNAQYRQGEARFLPKDLPSLCTHLIYAFAGMTNHLSTTEWNETLYQEFNG 61
QY 62 LKNKNSQLKTLAIGGNWFTAPFTAMVSTPENRQFTTTSVIKFLRQYFDFGLDFDWEYP 121
DB 62 LKKNMPLKTLAIGGNWFTQKFTDMVATANNRQTFVNSAIRFLRKYSFDFGLDLDWEYP 121
QY 122 GSRGSPPODKHLFTVLVOEMREAFQEAQKINPKRLMVTAAVAAGISNIQSGEYEPQLSQ 181
DB 122 GSQGSFAVDKERFTTLVQDLANAFQGEAQTSGKERLLLSAAVPAQGTVDAGYEVDKIAQ 181
QY 182 YLDYIHVMYDYLHGSWEGYTGENSEPLYKYPTDGTGNSAYLNVDYVMYKNDGAPAEKLI 241
DB 182 NLDVFNLMAYDFHGSWEKVTGHNSPLYKQESGAAASLNVDAAVQOMLQKGTGPASKLIL 241
QY 242 GFTYGHNFILSNPNTGIGAPTSAGPAGPYAKESGIWAYEICTFLKNGATQGWDAPO 301
DB 242 GMPTYGRSFTLASSSDTRVGAPATGSGTGPFTKEGMLAYEVCWS--KGATKQRIQDQ 299

Db 242 GMPYGRSFTLASSDTRVGAPATGSGTGPFTKEGMLAYEVCW--KGATKORI0DQ 299
Qy 302 EYPYAYQGNVWGYDNIXSFDIKAOWLKHNFKGAMVWALDLDFTGTCNQGKPELIST 361
Db 300 KVPYIFRDNQWVGFDVSEFVKVSYLKQKGLGGAMVWALDLDLDDPAGPSCNQGRIPIQT 359
Qy 362 LKKALGL 368
Db 360 LQELSL 366

Search completed: June 29, 2003, 21:03:29
Job time : 11.7788 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: June 29, 2003, 21:03:33 ; Search time 2957.5 Seconds
(without alignments)
4477.355 Million cell updates/sec

Title: US-10-004-219B-14
Perfect score: 2506
Sequence: 1 YQLTCYFTNWAQYRPLGRF.....QQNCOAGLVFDTSCDCNWA 455

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -MODEL=frame+ p2n.model -DEV=xl
-Q/cgn2_1/USPTO_spool/US10004219/runat_24062003_160228_8159/app_query.fasta_1.1877
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10004219 @CGN 1 1 6828 @runat_24062003_160228_8159 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=10 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sv.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sv.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2506	100.0	1625	9	AF290004 Homo sapi
2	2137	85.3	1369	6	AX405989 Sequence
3	2119	84.6	1354	9	AB025008 Homo sapi
4	2097.5	83.7	1526	10	AF154571 Mus muscu
5	2097.5	83.7	1538	10	BC011134 Mus muscu
6	2097.5	83.7	1557	10	BC034548 Mus muscu
7	2089.5	83.4	1530	10	AF290003 Mus muscu
8	2039	81.4	1529	4	AB051629 Bos tauru
9	1768.5	70.6	1188	9	AB025009 Homo sapi
10	1742	69.5	4250	9	AK098814 Homo sapi
11	1415	56.5	1525	10	M94584 Mus musculu
12	1407	56.1	1506	10	D87757 Mus musculu
13	1384	55.2	1209	10	AY049765 Mus muscu
14	1379	55.0	1209	10	AY065557 Mus muscu
15	1362.5	54.4	91569	9	AL356387 Human DNA
16	1353	54.0	93740	9	AL513202 Human DNA
17	1340.5	53.5	1562	5	BJA345054 Bufo japo
18	1323	52.8	1633	9	HSU29615 Human chito
19	1323	52.8	1636	6	AR206041 Sequence
20	1323	52.8	1636	6	AR212138 Sequence
21	1323	52.8	1636	6	AX108750 Sequence
22	1323	52.8	1643	6	AR172107 Sequence
23	1317	52.6	1656	6	AR206042 Sequence
24	1317	52.6	1656	6	AR212139 Sequence
25	1317	52.6	1656	6	AR212139 Sequence
26	1311.5	52.3	1710	9	HSU62662 Sequence
27	1311.5	52.3	1713	6	AR172108 Sequence
28	1225.5	48.9	1599	9	AK055165 Homo sapi
29	1011.5	40.4	2504	6	E09046 cDNA encodi
30	1011.5	40.4	2525	10	MUSO3P Mouse mRNa
31	1010.5	40.3	6439	3	AB074977 Haemaphys
32	985.5	39.3	1741	9	HUMH3G Human glyco
33	983.5	39.2	1801	9	BC008568 Homo sapi
34	983	39.2	1281	4	AY081150 Capra hir
35	980	39.1	3946	3	AY051988 Drosophil
36	972	38.8	1733	4	SSU19300 Sus scrofa
37	972	38.8	1733	4	SSGFP38RD S.scrofa 38
38	971.5	38.8	1149	6	E01500 cDNA encodi
39	971.5	38.8	1149	6	E01501 cDNA encodi
40	970.5	38.7	1418	9	HSU49835 Human YKL-3
41	970.5	38.7	1433	6	AR042834 Sequence
42	970.5	38.7	1434	9	HSU58314 Human chiti
43	970.5	38.7	1500	9	HSU58515 Human chiti
44	970.5	38.7	1526	6	AR042835 Sequence
45	966.5	38.6	1449	9	BC011460 Homo sapi

ALIGNMENTS

RESULT 1

AF290004 1625 bp mRNA linear PRI 27-FEB-2001
 LOCUS Homo sapiens acidic mammalian chitinase precursor, mRNA, complete
 DEFINITION cds.
 ACCESSION AF290004
 VERSION AF290004.1 GI:12597292
 KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1625)
 AUTHORS Boot, R.G., Blommaert, E.F., Swart, E., Chauharali-van der Vlugt, K.,
 Bijl, N., Moe, C., Place, A. and Aerts, J.M.
 TITLE Identification of a novel acidic mammalian chitinase distinct from
 chitotriosidase
 J. Biol. Chem. 276 (9), 6770-6778 (2001)
 JOURNAL 21125893
 MEDLINE 11085997
 PUBMED
 REFERENCE 2 (bases 1 to 1625)
 AUTHORS Boot, R.G., Verhoeck, M., Swart, E. and Aerts, J.M.
 TITLE Direct Submission
 JOURNAL Submitted (26-JUL-2000) Dept. of Biochemistry, Academic Medical
 Center, University of Amsterdam, Meibergdreef 15, Amsterdam 1105
 AZ, The Netherlands
 FEATURES
 Location/Qualifiers
 1..1625
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /tissue_type="stomach; lung"
 104..1534
 /EC_number="3.2.1.14"
 /note="chitinase precursor; AMCase; chitinase activity
 demonstrated by expression in COS cells; similar to
 sequences presented in GenBank Accession Numbers AB025008
 and AB025009"
 /codon_start=1
 /product="acidic mammalian chitinase precursor"
 /protein_id="AAAG60019.1"
 /db_xref="GI:12597293"
 /translation="MTKLLILTLGLVLILQLGSAVLQCYFTNNWAQYRGLGRFMPD
 NIDPCLTLHLIYAFAGQNNETIEWNVTLTYQAFNGLNKNSQLKTLAIGWNFG
 TAPFTAMVSTPNROFTTSVIFKFLQYDFDGLDFWEPGSRGSPQDKHLFTVLVQ
 ENREAEQAKNLKRLAMVTAAGAIGNISQYIPLDSVLDYIHVMYVDLHGSW
 EGYTGNISPLYKPTDTSNAYLVNVDVNNYKNGAPAEKLIIVGPTTYGHNFIILNP
 SNTGIGASTSGAGPAGYAKESGIWAYVEI CTFLKNGATQGDWAQEVYPAVQGNVW
 GDNIKSFIDKQWLKHNKFGAMWALIDDDFTGFCNQGKPLFLISTLUKALGLQSA
 SCTAPAPRIEPTTAAPSGSGNGSGSSSGSGSGSGGFCFCAVRANGLYPVANNRNFHC
 VNGVTYQQNCQAGLVFDTSDCCNWA"
 sig_peptide 104..166
 BASE COUNT 387 a 468 c 399 g 371 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 3,03e-184 Length: 1625
 Score: 2506.00 Matches: 455
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0
 US-10-004-219b-14 (1-455) x AF290004 (1-1625)
 QY 1 TyrlGluLeuThrCysTyrPheThrAsnTyrAlaGlnTyrArgProGlyLeuGlyArgPhe 20
 Db 167 TACCAGCTGACATGCTACTTACCACTAGTGGCCCGACGTACCGGCAGGCGCTGGGGCGCTTC 226
 QY 21 MetProAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArg 40
 Db 227 ATGCCTGACACATCGACCCCTGCTCTGTACCCACTGATCTACGCCCTTGTCTGGGAGG 286
 QY 41 GlnAsnGluIleThrThrIleGluTyrAsnAspValThrLeuTyrGlnAlaPheAsn 60

287 CAGAAACACGAGATCACCACCATCGAATGGAACGATGTGACTCTCTTACCAAGCTTTCAAT 346
 QY 61 GlyLeuLysAsnLysAsnSerGlnLeuLysThrLeuAlaIleGlyGlyTrpAsnPhe 80
 Db 347 GGCCTGAAATAAAGAACAGACGAGCTGAAATCTCTCGGCCATTGGAGGCTGGAACTTC 406
 QY 81 GlyThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThr 100
 Db 407 GGGACTGCCCTTTCACTGCCATGGTTTCTACTCTGAGAACCAGCAGACTTTTCATCAC 466
 QY 101 SerValIleLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyr 120
 Db 467 TCAGTCATCAATCTCTCGCCAGTATGAGTTTGAAGGCTGGACTTTGACTGGGAGTAC 526
 QY 121 ProGlySerArgGlySerProGlnAspLysHisLeuPheThrValLeuValGlnGlu 140
 Db 527 CCTGGCTCTCGTGGGAGGCCCTCTCAGCAACAGCATCTTCTACTGTCTGTGTGGAGGAA 586
 QY 141 MetArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThr 160
 Db 587 ATGCGTGAAGCTTTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 646
 QY 161 AlaAlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSer 180
 Db 647 GCTGCAGTAGCTGCTGCATCTCCAATATCCAGTCTGGCTATGAGATCCCACTGCA 706
 QY 181 GlnTyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr 200
 Db 707 CAGTACCTGGAGTATCATCTCATGTATGACCTACGACCTCCATGGCTCTCTGGGAGGCTAC 766
 QY 201 ThrGlyGluAsnSerProLeuTyrLysTyrProThrAspThrGlySerAsnAlaTyrLeu 220
 Db 767 ACTGGAGAGAACAGACCCCTCTTACAATATCCGAGTACACCCGCGCAGCAACGCTTACCTC 826
 QY 221 AsnValAspTyrValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeuIle 240
 Db 827 AATGTGGATTATGTGATGAACCTACTCGAAGGAGCAATGGAGCACCAGCTGAGAAGCTCATC 886
 QY 241 ValGlyPheProThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIle 260
 Db 887 GTTGGATTCCTTACCTATGGACACAACTTCATCTGGAGCAACCCCTCCAACTGGAATT 946
 QY 261 GlyAlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyIleTrp 280
 Db 947 GGTGCCCCACCTCTGTGTGCTGTCTCTGTGGCCCTATGCCAAGAGTCTGGGATCTGG 1006
 QY 281 AlaTyrTyrGluIleCysThrPheLeuLysAsnGlyAlaThrGlnGlyTrpAspAlaPro 300
 Db 1007 GCTTACTACGAGATCTGTACTCTTCTGAAAAATGGAGCCACTCAGGGATGGGATGCCCT 1066
 QY 301 GlnGluValProTyrAlaTyrGlnGlyAsnValTrpValGlyTyrAspAsnIleLysSer 320
 Db 1067 CAGGAAGTGCCCTTATGCTTATCAGGGCAATGTGGGTGGCTATGACAAACATCAAGAGC 1126
 QY 321 PheAspIleLysAlaGlnTrpLeuLysHisAsnLysPheGlyGlyAlaMetValTrpAla 340
 Db 1127 TTCGATATAAGGCTCAATGGCTTAAGCACAACAATTTGGAGGCGCCATGCTGGGCC 1186
 QY 341 IleAspLeuAspAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSer 360
 Db 1187 ATTGATCTGGATGACTTCACTGGCACTTTCTGCAACACAGGGCAAGTTTCCCCATACTCC 1246
 QY 361 ThrLeuLysLysAlaLeuGlyLeuGlnSerAlaSerCysThrAlaProAlaGlnProIle 380
 Db 1247 ACCCTGAAGAGGCCCTCGGCTTCAGAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGC 1306
 QY 381 GluProIleThrAlaAlaProSerGlySerGlyAsnGlySerGlySerSerSerSerGly 400
 Db 1307 GAGCCAATAACTGCTGCTCCAGTGCAGCGGGAACGGAGGCGGAGTAGCAGCTCTGGA 1366
 QY 401 GlySerSerGlyGlySerGlyPheCysAlaValArgAlaAsnGlyLeuTyrProValAla 420
 Db 1367 GGCAGCTCGGAGGAGTGGATTCTGT 1426

```
QY 421 AenAsnAlaGAsnAlaPheThrHisCysValAsnGlyValThrTyrGlnGlnAsnCysGln 440
Db 1427 AATAACAGAAATGCTCTGGACACTGCGTGAATGAGTCACTACAGAGAACTGCCAG 1486
QY 441 AlaGlyLeuValPheAspThrSerCysAspCysCysAsnTrpAla 455
Db 1487 GCGGGCTGTCTTCGACACCACTGCTGATTGCTGCAACTGGCA 1531

RESULT 2
AX405989 1369 bp DNA linear PAT 14-JUN-2002
LOCUS Sequence 404 from Patent WO0222660.
DEFINITION AX405989
ACCESSION AX405989
VERSION AX405989.1 GI:21439417
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Tang, Y.T., Liu, C., Zhou, P., Asundi, V., Zhang, J., Zhao, Q.A., Ren, F.,
Xue, A.J., Tang, Y., Wehrman, T. and Drmanac, R.T.
TITLE Novel nucleic acids and polypeptides
JOURNAL Patent: WO 0222660-A 404 21-MAR-2002;
HYSBO, INC. (US)
FEATURES
source 1..1369
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
154..1260
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD35096.1"
/db_xref="GI:21439418"
/translation="MVSTPENRQTFITSVIKELROYFDGLDFDWEYPGSRGSPQDK
HLFTVLQSMRAFQEAQKINKPRLMTVAAGISNIQSGYEIPQLSQYLDYIHVM
TYDLHWSGEGTYGSPLYKYPTDTGSNAYLVNDVYMNWKONGAPAEKLI VGFPTYG
HNFILSNSTGCTGSPAGPVAKESGLWAYEICTFLKNGATCGMDAPQEVY
AYQGNVWGYDNIKSFIDIKAWLKNKFGAMWALDLDFTGTCNOKGKFLISTLK
KALGQASACTAPADIEPITAPSGSGSGSSGSSGSGSGSGSGSGSGSGSGSGSGSG
NRNATFWHCYNGVTYQQNCOAGLYVFTDSCDCNNA"
BASE COUNT 327 a 393 c 341 g 308 t
ORIGIN

Alignment Scores:
Pred. No.: 7.17e-156 Length: 1369
Score: 2137.00 Matches: 390
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.74% Mismatches: 0
Query Match: 85.28% Indels: 0
DB: 6 Gaps: 0

US-10-004-219b-14 (1-455) x AX405989 (1-1369)
QY 65 LysAsnSerGlnLeuLysThrLeuLeuAlaLeuGlyTyrAsnPheGlyThrAlaPro 84
Db 85 CGGAACAGCAGCTGAAACTCTCTCGCCCATTTGGAGCTGGAACTTCGGACTGCCCT 144
QY 85 PheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThrSerValIleLys 104
Db 145 TTCACCTGCCATGTTCTACTCTCTGAGAACCGCCAGACTTTCATCACTCACTCAAAA 204
QY 105 PheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyrProGlySerArg 124
Db 205 TTCCTCGCCAGTATGATGTTGACGGCTGAGCTTTGACTGGGAGTACCCCTGCTCGT 264
QY 125 GlySerProGlnAspLysHisLeuPheThrValLeuValGlnGluMetArgGluAla 144
Db 265 GGGAGCCCTCTCAGGACAGCATCTCTTCACTGTCTGTGTGAGGAAATGCGTGAAGCT 324
QY 145 PheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThrAlaValAla 164
```

```
Db 325 TTTGAGCAGGAGCCAGCAGATCAACAGCCAGCTGATGGTCACCTGCTGCATAGCT 384
QY 165 AlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSerGlnTyrLeuAsp 184
Db 385 GCTGGCATCTCAATATCCAGTCTGGCTATGAGATCCCCCACTCTCACAGTACCTGGAC 444
QY 185 TyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThrGlyGluAsn 204
Db 445 TACATCCATGTTCATGACCTACGACCTCCATGGCTCTGGAGGGCTACACTCGAGAGAAC 504
QY 205 SerProLeuTyrLysTyrProThrAspThrGlySerAsnAlaTyrLeuAsnValAspTyr 224
Db 505 AGCCCCCTCTCAAAATACCGACTCACCGGACCAACCGCTACTCTCAATGTGATAT 564
QY 225 ValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeuIleValGlyPhePro 244
Db 565 GTCATGAATCTACTGGAGGACAAATGGAGCACAGCTGAGAGCTCATCTGTGGATTCCCT 624
QY 245 ThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIleGlyAlaProThr 264
Db 625 ACCTATGGACACAACTTCATCTCTGAGCAACCCCTCCACACTGGAATTGGTCCCCACC 684
QY 265 SerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyIleTrpAlaTyrTyrGlu 284
Db 685 TCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 744
QY 285 IleCysThrPheLeuLysAsnGlyAlaThrGlnGlyTyrAspAlaProGlnGluValPro 304
Db 745 ATCTGTACTCTTCTGAAATAATGGAGCCACTCAGGATGGGATGCCCTCAGAGAGTGCCT 804
QY 305 TyrAlaTyrGlnGlyAsnValTrpValGlyTyrAspAsnIleLysSerPheAspIleLys 324
Db 805 TATGCTTATCAGGCAATGTGTGGTGGCTGCTATGCAACATCAAGAGCTTCGATATTAAG 864
QY 325 AlaGlnTrpLeuLysHisAsnLysPheGlyGlyAlaMetValTrpAlaIleAspLeuAsp 344
Db 865 GCTCAATGGCTTAAAGCAACAAATAATGGAGCGGCATGCTGTCGGCCATTGATCTGGAT 924
QY 345 AspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSerThrLeuLysLys 364
Db 925 GACTTCACTGGCACATTTCTGCAACCAAGGCAAGTTTCCCTAATCTCCACCTCTGAAGAAG 984
QY 365 AlaLeuGlyLeuGlnSerAlaSerCysThrAlaProAlaGlnProIleGluProIleThr 384
Db 985 GCGCTCGGCTCGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGT 1044
QY 385 AlaAlaProSerGlySerGlyAsnGlySerGlySerSerSerSerSerSerSerSerGly 404
Db 1045 GCTGCTCCAGTGGCAGCGGGAACGGGAGCGGAGTAGCAGCTCTGGAGGACAGCTCGGGA 1104
QY 405 GlySerGlyPheCysAlaValAlaArgAlaAsnGlyLeuTyrProValAlaAsnAsnArg 424
Db 1105 GGCAGTGGATTCTGTGCTGTGAGCAACAGGCTCTACCCCTGGTGCATTAACAGAAAT 1164
QY 425 AlaPheTrpHisCysValAsnGlyValThrTyrGlnGlnAsnCysGlnAlaGlyLeuVal 444
Db 1165 GCTTCTGGCACTGCTGATGAGTCACTACAGCAGACAGTGCAGGAGCGGCTGTGTC 1224
QY 445 PheAspThrSerCysAspCysCysAsnTrpAla 455
Db 1225 TTCGACACCACTGTGATTGCTGCAACTGGCA 1257

RESULT 3
AB025008 1354 bp mRNA linear PRI 25-NOV-1999
LOCUS Homo sapiens TSA1902-L mRNA for novel member of chitinase family,
DEFINITION complete cds.
ACCESSION AB025008
VERSION AB025008.1 GI:6467176
KEYWORDS novel member of chitinase family; TSA1902-L.
SOURCE Homo sapiens cDNA to mRNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

Mammalia; Euthera; Primates; Catarrhini; Homidae; Homo.

REFERENCE

1 (eites)

AUTHORS

Saito, A., Ozaki, K., Fujiwara, T., Nakamura, Y. and Tanigami, A.

TITLE

Isolation and mapping of a human lung-specific gene, TSA1902,

JOURNAL

encoding a novel chitinase family member

MEDLINE

Gene 239 (2), 325-331 (1999)

REFERENCE

20018184

TITLE

(bases 1 to 1354)

AUTHORS

Saito, A., Ozaki, K., Fujiwara, T., Takahashi, E. and Tanigami, A.

JOURNAL

Submitted (16-MAR-1999) Akihiko Saito, Otsuka Pharmaceutical Co.,

TITLE

Ltd., Otsuka GEN Research Institute; 463-10 Kagasuno Kawauchi-cho,

JOURNAL

Tokushima, Tokushima 771-0192, Japan [E-mail: saito@otsuka.gr.jp,

FEATURES

Tel: 81-88-2888, Fax: 81-88-637-1035]

source

Location/Qualifiers

gene

1. .1354

CDS

/organism="Homo sapiens"

gene

/db_xref="taxon:9606"

CDS

1. .1354

gene

/genes="TSA1902-L"

CDS

157..1263

gene

/genes="TSA1902-L"

CDS

/codon_start=1

gene

/product="novel member of chitinase family"

CDS

/protein_id="BAA86980.1"

gene

/db_xref="GI:6467177"

CDS

/translation="MUSTPENRQTFTTSVKEFLRVEFDGLDFDWEYPCSGSPPODK

gene

HFTVLVQEMREAFEGEAKQINKPLMVTAAVAGISNTQSGVEIPQLSQYLDYHVM

CDS

TYDLHSGWEGYTGNSPLKYPTDGSNAYLVNVMYKONGAPAEKLIYGFPTYG

gene

HNFLSPNTGIGTAPSGAGPYAKESGWAYEICTFLKNGTQMDAPQEVY

CDS

ATGQNVGVNDVNPDIKAQMLKHNFGGAMVWIDLDFTFTFCNQKFLISTLK

gene

KALGLOSCTAPQAPQIEPTITAPSGSGNGSGSSGSGSGGFCAGRANGLYPVAN

CDS

NRNAFWHCYGVTVQNCQAGLVDFNSCDCCNWA"

BASE COUNT

311 a 387 c 348 g 308 t

ORIGIN

1.73e-154 Length: 1354

Alignment Scores:

2119.00 Matches: 387

Pred. No.:

99.49% Conservative: 1

Score:

99.23% Mismatches: 2

Percent Similarity:

84.56% Indels: 0

Best Local Similarity:

9 Gaps: 0

Query Match:

US-10-004-219B-14 (1-455) x AB025008 (1-1354)

DB:

66 AnSerGlnLeuLysThrLeuLeuAlaLeuGlyGlyTyrPheGlyThrAlaProphe 85

QY

91 AACAGCCAGCTGAAACCTCTCTGGCCATTGGAGGCTGGAACCTTCAGGACTGCCCTTTTC 150

DB

86 ThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThrSerValIleLysPhe 105

QY

151 ACTGCCATGTTCTACTCTGAGAACCCGACACTTTCATCACTCAGTCATCAAAATTC 210

DB

106 LeuArgGlnTyrGluPheAspGlyLeuAspPheAspTyrPheGlySerArgGly 125

QY

211 CTGGCCAGTATGAGTTTACGGGCTGGACTTTGATGGGAGTACCTTGGCTCTCTGGG 270

DB

126 SerProProGlnAspLysHisLeuPheThrValLeuValGlnGluMetArgGluAlaPhe 145

QY

271 AGCCCTCCCTCAGGACCAAGCATCTCTTCACTGTCTGGTGGAGAAATGGCTGAAGCTTTT 330

DB

146 GluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThrAlaAlaValAla 165

QY

331 GACGAGAGGCCAAGCAGATCAACAGCCAGGCTGATGCTACTGCTGAGTAGTGTCT 390

DB

166 GlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSerGlnTyrLeuAspTyr 185

QY

391 GGCACTCTCAATATCCAGCTGGCTGATGATGATGATGATGATGATGATGATGATGATGAT 450

DB

186 IleHisValMetThrTyrAspLeuHisGlySerTyrGluGlyTyrThrGlyGluAsnSer 205

QY

451 ATCCATGTATGACCTACGACCTCCATGCTCTCTGGAGGCTTACACTGGAGAGAACAGC 510

DB

206 ProLeuTyrLysTyrProThrAspThrGlySerAsnAlaTyrLeuAsnValAspTyrVal 225

QY

511 CCCCTCTCAATATCCAGCTGACACCGGCGAGCAAGCTTACTCTCACTGGATGATGTC 570

DB

226 MetAsnTyrTyrPheAspAsnGlyAlaProAlaGluLysLeuIleValGlyPheProThr 245

QY

571 ATGAACTACTGGAAGGCAATGAGGACCACTGAGAGCTCATGCTTGGATTCCTTACC 630

DB

246 TyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIleGlyAlaProThrSer 265

QY

631 TATGACACCACTTCTCTGAGCAACCCCTCCACACTGGAATTTGGTCCCCCCTCTCT 690

DB

266 GlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyIleTyrAlaTyrTyrGluIle 285

QY

691 GGTGCTGGTCTCTGGGCCCTATGCCAAGAGTCTGGGATCTGGGCTTACTACGAGATC 750

DB

286 CysThrPheLeuLysAsnGlyAlaThrGlnGlyTyrAspAlaProGlnGluValProTyr 305

QY

751 TGTACTCTTCTGAAATGAGGCACTCAGGATGGGATGGCCCTCAGGAAGTGCCTTAT 810

DB

306 AlaTyrGlnGlyAsnValTyrValGlyTyrAspAsnIleLysSerPheAspIleLysAla 325

QY

811 GCTATATCAGGGCAATGTGTGGTGGCTATGACACGTCACAGCTTCGATATTAGGCT 870

DB

326 GlnTyrLeuLysHisAsnLysPheGlyGlyAlaMetValTyrAlaIleAspLeuAspAsp 345

QY

871 CAATGGCTTAAGCAACAAATTTGAGGCGCATGGCTTGGGCCATTTGATCTGGATGAC 930

DB

346 PheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSerThrLeuLysLysAla 365

QY

931 TTCACTGGCACTTTCTGCAACCCAGGCAAGTTTCCCTTAATCTCCACCTGGAAGAGGC 990

DB

366 LeuGlyLeuGlnSerAlaSerCysThrAlaProAlaGlnProIleGluProIleThrAla 385

QY

991 CTGGCCCTGCAGAGTGCAGTTGACGGCTCAGCTCAGCTCAGCCATTCAGCAATACTGCT 1050

DB

386 AlaProSerGlySerGlyAsnGlySerGlySerSerSerSerSerGlySerGlyGly 405

QY

1051 GCTCCAGCTGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1110

DB

406 SerGlyPheCysAlaValAlaArgAlaSerGlyLeuTyrProValAlaAsnAsnArgAsnAla 425

QY

1111 AGTGGATTCTGTCTGGCAGAGCAACCGGCTCTACCCCGTGGCAAAATACAGAAATGCC 1170

DB

426 PheThrHisCysValAsnGlyValThrTyrGlnGlnAsnCysGlnAlaGlyLeuValPhe 445

QY

1171 TTCTGGCACTGCTGTAATGGAGTACAGTACAGCAGAGAACTGCCAGCGCGGCTTGTCTTC 1230

DB

446 AspThrSerCysAspCysCysAsnTyrAla 455

QY

1231 GACACCACTGCTGATTTGTCGCACTGGGCA 1260

DB

AF154571 1526 bp mRNA linear ROD 08-FEB-2000

QY

Mus musculus putative chitinase precursor (YNL) mRNA, partial cds.

DB

AF154571 GI:6934189

QY

Mus musculus.

DB

Mus musculus

QY

Mus musculus

DB

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

QY

Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

DB

1 (bases 1 to 1526)

QY

YNL, A Putative Mouse Chitinase

DB

Unpublished

QY

2 (bases 1 to 1526)

DB

Price, P.A., Harris, S.C. and Williamson, M.K.

QY

Direct Submission

DB

Submitted (26-MAY-1999) Biology, University of California, San

QY

Diego, 9500 Gilman Drive, Mail Code 0368, La Jolla, CA 92093-03687,

DB

RESULT 4

QY

AF154571

DB

LOCUS

QY

DEFINITION

DB

ACCESSION

QY

VERSION

DB

KEYWORDS

QY

SOURCE

DB

ORGANISM

QY

REFERENCE

DB

AUTHORS

QY

TITLE

DB

JOURNAL

QY

REFERENCE

DB

AUTHORS

QY

TITLE

DB

JOURNAL

USA

FEATURES

source

Location/Qualifiers

1. .1526
 /organism="Mus musculus"
 /strain="ATCC1476668"
 /db_xref="ATCC:1476668"
 /db_xref="taxon:10090"
 /clone="IMAGE 1209227"
 /tissue_type="skin"
 <1. .1526
 /gene="YNL"
 <1. .1419
 /gene="YNL"
 /codon_start=1
 /product="putative chitinase precursor"
 /protein_id="AAF31644.1"
 /db_xref="GI:6914190"
 /translation="AKLLIVTGLALLNQLGSAYNLICYFTNWAQYRGLGSKFPDD
 INPCLTHLIYAFAGQNNIEITIEWNDVLYKAFNDLKNRNSKLKLLAIGWNFGT
 APFTMTVSTNSQRQFTITSVIFKRLQYGFGLDLWEPYSGSRGSPDKHLFTLVVKE
 MREAPEQAIENRPLMVAAGVIGNIOAGYEIPSLKYLDFHVMYTDLHGSWE
 GYTGENSLYKPTPTGSNAYLVNVDYVMYKNNNGAPAEKLVGPFYGHYTFILRNPS
 DNGIGAPSGDGPAGPYRQAGFWAYEICTFLRSGATEVWDQSEVPYAKANEWL
 YDNISFSVKAQMLKQNFSGGAWIWDLDLDFGSCDQGFPLSTLTKALGISTEG
 CTAPDVPEPVTPTPGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSG
 TYQHCQAGLVFDTSNCNCNP"
 <1. .60
 /gene="YNL"
 61. .1416
 /gene="YNL"
 /product="putative chitinase"
 394. .417
 /gene="YNL"
 /note="putative conserved region; active site"
 415. .417
 /gene="YNL"
 /note="putative invariant Glu; active site"
 383 a 381 c 364 g 398 t

sig_peptide

mat_peptide

misc_feature

misc_feature

BASE COUNT

ORIGIN

Alignment Scores:

Pred. No.: 9,14e-153 Length: 1526
 Score: 2097.50 Matches: 370
 Percent Similarity: 90.31% Conservative: 40
 Best Local Similarity: 81.50% Mismatches: 41
 Query Match: 83.70% Indels: 3
 DB: 10 Gaps: 1

US-10-004-219B-14 (1-455) x AF154571 (1-1526)

QY 1 TyrGlnLeuThrCysTyrPheThrAsnTTPAlaGlnTyrArgProGlyLeuGlyArgPhe 20
 DB 61 TACAATCTGATATGCTATTTTACCAACTGGGCCAGTATCGGCCAGGTCTGGGAGCTTC 120
 QY 21 MetProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArg 40
 DB 121 AAGCCTGATGACATTAAACCCCTGCTGTACTCACCTGATCTATGCTTGTGGATG 180
 QY 41 GlnAsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrGlnAlaPheAsn 60
 DB 181 CAGAACAAATGAGATACCAACCAATAGAAATGATGTTACTCTCTATAAAGCTTTCAAT 240
 QY 61 GlyLeuLeuAsnIleAsnSerGlnLeuLeuThrLeuLeuAlaIleGlyGlyTyrAsnPhe 80
 DB 241 GACTTGAAAACAGGAACAGCAACTGAAACCCCTCTCGCAATGGAGGCTGGAACTTT 300
 QY 81 GlyThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThr 100
 DB 301 GGAACCTGCTCTTCACTACCATGGTTTCCACTTCTCAGAACCGCCAGACCTTCATTACC 360
 QY 101 SerValIleLeuPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyr 120
 DB 361 TCAGTCATCAAAATTTCTGCGTCAGTATGGGTTTGATGGACTGGACCTGGACCTGGGAATAC 420

RESULT 5

BC011134

LOCUS

DEFINITION

BC011134 1538 bp mRNA linear ROD 07-AUG-2002
 Mus musculus, similar to eosinophil chemotactic cytokine, clone

QY 121 ProGlySerArgGlySerProProGlnAspIlyshisLeuPheThrValLeuValGlnGlu 140
 DB 421 CCAGGCTCAGCTGGGAGCCCTCCTCAGGACAAGCATCTCTTCACTGTCTGTGTGAAGGAA 480
 QY 141 MetArgGluAlaPheGluGlnGlnAlaIlyshisGlnIleAsnLysProArgLeuMetValThr 160
 DB 481 ATGCGTGAAGCTTTTGGACGAGGCTATTGAGAGCAACAGCCAGAGCTGATGGTTACT 540
 QY 161 AlaAlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSer 180
 DB 541 GCTGCTGATAGCTGGTGGGATTTCCAAATCCAGGCTGCTGATGAGATCCCTGAGATTCT 600
 QY 181 GlnTyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr 200
 DB 601 AAGTACCTGGATTTTCATCATGTCATGATGATGATGATGATGATGATGATGATGATGAT 660
 QY 201 ThrGlyGluAsnSerProLeuTyrIlyshisTyrProThrAspThrGlySerAsnAlaTyrLeu 220
 DB 661 ACTGGGAGAAATAGTCTCTTTTACAAATACCTACTAGAGCTGGTAGCAATGCCTACCTC 720
 QY 221 AsnValAspTyrValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeuIle 240
 DB 721 AATGTTGGATTATGTCATGAATTTTGGAGAAACAATGGAGCCCACTGAGAGCTGAGAGCTCAT 780
 QY 241 ValGlyPheProThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIle 260
 DB 781 GTTGGATTTCCAGAGTATGGACACACCTTTCATCTCGAGAAACCCCTCTGATAATGGAAT 840
 QY 261 GlyAlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyIleTrp 280
 DB 841 GGTGCCCTTACCTCTGGTGTGGTGGCTCTGTGGGCTCTATACCCAGCAGAGCTGGTCTGG 900
 QY 281 AlaTyrTyrGluIleCysThrPheLeuLysAsnGlyAlaThrGlnGlyTrpAspAlaPro 300
 DB 901 GCCTACTATGAGATTGACCTTCTGAGAGTGGAGCCACTGAGCTTGGATGCTGCTCC 960
 QY 301 GlnGluValProTyrAlaTyrGlnGlyAsnValTrpValGlyTyrAspAsnIleLysSer 320
 DB 961 CAAAGAAGTGCCCTATGCTATAAGGCCAACAGTGGCTTGGCTATGACATAATCAAGAGC 1020
 QY 321 PheAspIleLysAlaGlnTrpLeuLysHisAsnLysPheGlyGlyAlaMetValTrpAla 340
 DB 1021 TTCAGTGTAGGCTCAGTGGCTTAAGCAGACAAATTTTGGAGGTGCCATGATGCTGGGCC 1080
 QY 341 IleAspLeuAspAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSer 360
 DB 1081 ATTGACCTTGATGACTTCACTGGCTCTTCTGTGATCAGGGAAATTTCTCTGACTTCT 1140
 QY 361 ThrLeuLysLysAlaLeuGlyLeuGlnSerAlaSerCysThrAlaProAlaGlnProIle 380
 DB 1141 ACTTTGAACAAAGCCCTTGGCATATCCACTGAAGGTTCACACAGCTCCTGACGTGCTTCC 1200
 QY 381 GluProIleThrAlaAlaProSerGlySerGlyAsnGlySerGlySerSerSerGly 400
 DB 1201 GAGCCAGTGACTACTCTCTCCA-----GGAAGTGGAGTGGGGGTGGAAAGCTCCGGA 1251
 QY 401 GlySerSerGlyGlySerGlyPheCysAlaValAlaGlnAsnGlyLeuTyrProValAla 420
 DB 1252 GGAAGCTCTGGAGGAGTGGATTTCTGTCCGCAAAAGCAGATGGGCTCTTACCTGTGGCA 1311
 QY 421 AsnAsnArgAsnAlaPheTrpHisCysValAsnGlyValThrTyrGlnGlnAsnCysGln 440
 DB 1312 GATGACAGAAATGCTTTTGGCAGTGCATCATGAAATCACAATACAGCAGCATGTGCA 1371
 QY 441 AlaGlyLeuValPheAspThrSerCysAspCysCysAsnTrp 454
 DB 1372 GCAGGGCTGTGTTTGTATACCAGCTGTAATGCTGCAACTGG 1413

MGC:18771 IMAGE:4165150, mRNA, complete cds.
 BC011134
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 house mouse:
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 1538)
 Strausberg, R.
 Direct Submission
 Submitted (25-JUL-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapsb@mail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: amg@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
 Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
 Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 24 Row: e Column: 11.

FEATURES

Location/Qualifiers
 i. 1538
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /map="FVB/N"
 /clone="MGC:18771 IMAGE:4165150"
 /tissue_type="Salivary gland, 10 week old female mouse"
 /clone_lib="NCI CGAP_SG2"
 /lab_host="DH10B"
 /notes="Vector: pCMV-SPORT6"
 343..1420
 /codon_start=1
 /product="Similar to eosinophil chemotactic cytokine"
 /protein_id="AAH1134.1"
 /db_xref="GI:15029822"
 /translation="MVSTQNRQTFTTSVIFKRLQYFGDLDDWEYPSGRSPQDK
 HLTFLVEMRFAFEQAEISNRPRLMVTAAGGISNTQAGVEIPELSKYLDFFHVM
 HTDLSGWSGYTGENSEPLVKYPTETGSNAYLVNVMYKNGNAPAEKLIYGFPEYG
 HTFLNPSDNGIGTSGDGPAGPYTROAGFWAYVEICTFLRSQATEVMDASQEVY
 AKANRWLGDNIKGSFVZAQWLKQNNFGNMIWALDDFTFGSCDQCKFELSTLN
 KALISTECTADVPDSEPVPTTPPGSGSGSGSGSGSGSGSGGFCADKADGLYPVADURN
 AFQWCTINGITYQHQAGLGVFDTSNCNWP"
 AFQWCTINGITYQHQAGLGVFDTSNCNWP"
 BASE COUNT 393 a 382 c 365 g 398 t
 ORIGIN

Alignment Scores:
 Pred. No.: 9.23e-153 Length: 1538
 Score: 2097.50 Matches: 370
 Percent Similarity: 90.31% Conservative: 40
 Best Local Similarity: 81.50% Mismatches: 41
 Query Match: 83.70% Indels: 3
 DB: 10 Gaps: 1

US-10-004-219b-14 (1-455) x BC011134 (1-1538)

QY 1 TyrGlnLeuThrCysTyrPheThrAsnTyrAlaGlnTyrArgProGlyLeuGlyArgPhe 20
 DB 62 TACAATCTGATGCTATTTTACCACTATCGGCCCAAGTATCGGCCAGGTCTGGGAGCTTC 121
 QY 21 MetProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArg 40

122 AAGCCTGATGACATTAACCCCTGCTGTACTACCTGATATGCTGCTGCGATG 181
 QY 41 GlnAsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrGlnAlaPheAsn 60
 DB 182 CAGAACATGAGATCACCACCATAGATGAATGATGTTACTCTCTATAAAGCTTTCAAT 241
 QY 61 GlyLeuLysAsnLysAsnSerGlnLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPhe 80
 DB 242 GACTTGAAAAACAGGAACAGCAAACTCTCTGGCAATTTGGAGGCTGGAATTT 301
 QY 81 GlyThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThr 100
 DB 302 GGAATCTGCTCTTCTCACTACCATGCTTTCCACTTCTCAGAACCCGACACCTTCATTACC 361
 QY 101 SerValIleLeuPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyr 120
 DB 362 TCAGTCATCAAAATTTCTCGTCAGTATGGTTTGTAGTACCTGGACTGGACTGGGAATAC 421
 QY 121 ProGlySerArgGlySerProGlnAspLysHisLeuPheThrValLeuValGlnGlu 140
 DB 422 CCAGGCTCAGCTGGGAGCCCTCTCAGGACAGCATCTCTTCACTGTCTCTGTGTAAGGAA 481
 QY 141 MetArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThr 160
 DB 482 ATGCGTGAAGCTTTTTCAGCAGGAGGCTATTGAGAGCAACAGGCCACAGCTGATGTTACT 541
 QY 161 AlaAlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSer 180
 DB 542 GCTGCTGTAGCTGGTGGGATTTCCAACTCCAGGCTGGCTATGAGATCCCTGAACTTCT 601
 QY 181 GlnTyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr 200
 DB 602 AAGTACCTGGATTTTCACTCACTGATGATGATGATGATGATGATGATGATGATGATGAT 661
 QY 201 ThrGlyGluAsnSerProLeuTyrIleThrTyrProThrAspThrGlySerAsnAlaTyrLeu 220
 DB 662 ACTGGGAGAAATAGTCTCTTTTACAAATACCTCTAGACTGGTGGTGGTGGTGGTGGTGGT 721
 QY 221 AsnValAspTyrValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeuIle 240
 DB 722 AATGTGGATTATGTATGAACCTATTGGAGAACAACTAGGAGCCCGAGCTGAGAGTCAAT 781
 QY 241 ValGlyPheProThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIle 260
 DB 782 GTTGGATTTCCAGAGATATGGACACACCTTCTCTGAGAAACCCCTCTGATAATGGAAT 841
 QY 261 GlyAlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyIleTrp 280
 DB 842 GGTGCCCTTACCTCTGTGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 901
 QY 281 AlaTyrTyrGluIleCysThrPheLeuLysAsnGlyAlaThrGlnGlyTrpAspAlaPro 300
 DB 902 GCCTACTATGAGTTGACCTTTCTGAGAGTGGAGCCAGCTGAGGTCTGGATGCTGCTCC 961
 QY 301 GlnGluValProTyrAlaTyrGlnGlyAsnValTrpValGlyTyrAspAsnIleLysSer 320
 DB 962 CAAGAAGTGGCTTATGCTTAAAGCCCAAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1021
 QY 321 PheAspIleLysAlaGlnTrpLeuLysHisAsnLysPheGlyGlyAlaMetValTrpAla 340
 DB 1022 TTCACTGTTAAGGCTCAGTGGCTTAAAGCAGAACAAATTTGGAGGTCGCCATGATCTGGGCC 1081
 QY 341 IleAspLeuAspAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSer 360
 DB 1082 ATTGACCTTGATGACTTCACTGGCTCTTCTGTGATCAGGGAATTTCTCTGACTCTCT 1141
 QY 361 ThrLeuLysLysAlaLeuGlyLeuGlnSerAlaSerCysThrAlaProAlaGlnProIle 380
 DB 1142 ACTTTGAACAAAGCCCTTGGCATATCCACTGAAGGTTGCACAGCTCTCTGACGTGCTTCC 1201
 QY 381 GluProIleThrAlaAlaProSerGlySerGlyAsnGlySerGlySerSerSerSerGly 400

Db 1202 GAGCCAGTGACTACTCTCTCA-----GGAGGTGGAGTGGGGTGGAGCTCCGGA 1252

Qy 401 GlySerSerGlyGlySerGlyPheCysAlaValArgAlaAsnGlyLeuTyrProValAla 420

Db 1253 GGAAGCTCTGGAGGAGTGTCTGTGGCGGACAAAGAGAGTGGCTCTACCTCTGTGGCA 1312

Qy 421 AsnAsnArgAsnAlaPheTrpHisCysValAsnGlyValThrTyrGlnGlnAsnCysGln 440

Db 1313 GATGACAGAAATGCTTTTGGCAGTGCATCAATGGAATCACATACCAGCAGCATGTGCA 1372

Qy 441 AlaGlyLeuValPheAspThrSerCysAspCysAsnTrp 454

Db 1373 GCAGGCTTGTGTTTATACACAGCTGTAATGCTGCACTGG 1414

RESULT 6

BC034548

LOCUS

DEFINITION Mus musculus, chitinase, acidic, clone MGC:19045 IMAGE:4189080, mRNA, complete cds.

ACCESSION BC034548

VERSION BC034548.1 GI:21961190

KEYWORDS MGC.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 1557)

AUTHORS Strausberg, R.

TITLE Direct Submission

JOURNAL Submitted (24-JUL-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT Contact: MGC help desk

Email: cgabs-remail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center

Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

Contact: angbcm.tmc.edu

Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M., Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAK Plate: 24 Row: h Column: 16

This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.

Location/Qualifiers

1. .1557

/organism="Mus musculus"

/db_xref="taxon:10090"

/map="FVB/N"

/clone="MGC:19045 IMAGE:4189080"

/tissue_type="Salivary gland, 10 week old female mouse"

/clone_lib="NCI CGAP_SG2"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

4. .1425

/codon_start=1

/product="chitinase, acidic"

/protein_id="AAH34548.1"

/db_xref="GI:21961191"

/db_xref="LocusID:81600"

/translation="WAKLLVTLGLALLNAQGSAYNICVPTNWAQYRPGIGSKFPD DNPCLCHLIYAFAGMKNNEITTEWVDVLYKAPNDLKNRNSKTLTLAGWNFG TAPFTTMVTSQNRQTFTSVIKFLRQGFGLDLWDVEYPSGRSPQDKHLFTVLVK

CDS

EMREAFQEAIESNRPLMVTAAVAGGISNIQAGYEIPELSKYLDIFIHVMTYDLHGWS

EGYTGNSPLKYKPTETGTSNAYLVNVDYVWYKNGNAPAEKLI VGFPEYGHFTILRNP

SDNGIGAPTSQDGPAGPTROAGPWVAYEICTFLRSGATEKVDWDAQSEVPYAYKANWL

GYDNITKSFVKAQWLKQNNFGGAMTALDLDFTGTFSCDQKKEPLTSLINKALGISTE

GCTAPFSPBVPITPPGGSGGGSGSGSGGSGGFCADKADGLYFVADDRNFAWQCING

ITYQHCQAGGLVDFSCNCCNWP"

BASE COUNT 410 a 382 c 366 g 399 t

ORIGIN

Alignment Scores:

Prod. No.: 9.37e-153 Length: 1557

Score: 2097.50 Matches: 370

Percent Similarity: 90.31% Conservative: 40

Best Local Similarity: 81.50% Mismatches: 41

Query Match: 83.70% Indels: 3

DB: 10 Gaps: 1

US-10-004-219B-14 (1-455) x BC034548 (1-1557)

Qy 1 TyrGlnLeuThrCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPhe 20

Db 67 TACAATCTGATATGCTATTTACCAACTGGGCCAGTATCGCCAGGCTCGGGAGCTTC 126

Qy 21 MetProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArg 40

Db 127 AAGCCTGATGACATAACCCCTGCTGTACTACCTGATCTATGCTCTTCTGGGATG 186

Qy 41 GlnAsnAsnGluIleThrThrIleGlnTrpAsnAspValThrLeuTyrGlnAlaPheAsn 60

Db 187 CAGAACATGATGATCACCACCATAGATGGAATGATGTTACTCTCTATAAAGCTTTCAAT 246

Qy 61 GlyLeuLysAsnLysAsnSerGlnLeuLysThrLeuAlaIleGlyGlyTrpAsnPhe 80

Db 247 GACTTGAATAACAGGACAGCAACTGAAACCTCTCTGGCAATTTGGAGCTTGGAACTTT 306

Qy 81 GlyThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThr 100

Db 307 GGAACCTGCTCTCTTCACTACCATGTTTCCACTTCTCAGAACCCGCGACACCTTCATTACC 366

Qy 101 SerValIleLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyr 120

Db 367 TCAGTCATCAAAATTTCTGCTGATGTTGTTGATGATGAGCTGGACTGGGAATAC 426

Qy 121 ProGlySerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValGlnGlu 140

Db 427 CCAGGCTCAGTGGAGGAGCCCTCTCAGCAACAGCATCTCTTCACTGTCTGCTGGTGAAGNA 486

Qy 141 MetArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThr 160

Db 487 ATGCGTGAAGCTTTTGAGCAGGAGCTATTGAGACCAACAGGCCCCAGACTGATGTTACT 546

Qy 161 AlaAlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSer 180

Db 547 GCTGCTGTAGCTGGTGGGATTTCCAACATCCAGGCTGGCTATGAGATCCCTGAACTTTCT 606

Qy 181 GlnTyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr 200

Db 607 AAGTACCTGGATTTTCATGTCATGTCATGATGATGATGATGATGATGATGATGATGATG 666

Qy 201 ThrGlyGluAsnSerProLeuTyrIleTyrThrAspThrGlySerAsnAlaTyrLeu 220

Db 667 ACTGGGGAGATAGTCTCTTTTACAATACTCTAGACTGGTGGTGGTGGTGGTGGTGGTGG 726

Qy 221 AsnValAspTyrValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeuIle 240

Db 727 AATGTGGATTATGTCATGAACCTATTGGGAAGAACAAATGGAGCCCGCCAGCTGAGAA 786

Qy 241 ValGlyPheProThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIle 260

Db 787 GTTGGATTTCCAGAGTAGTGGACACACCTTCATCTCGAAGAACCCCTCTGATTAATGAA 846

Qy 261 GlyAlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyIleTrp 280


```

Db      726 AATGTGGATTATGTCATGAAGCAATATGAGAGCCAGCTGAGAAGCTCAT 785
Qy      241 ValGlyPheProThrTyrrGlyHisAsnPhelLeuSerAsnProSerAsnThrGlyIle 260
Db      786 GTTGGATCCAGAGTATGAGACACACTTCATCTGAGAAACCCCTCTGATATGAAT 845
Qy      261 GlyAlaProThrsertGlyAlaGlyProAlaGlyProTyrrAlaLysGluSerGlyIleTrp 280
Db      846 GGTGCCCTTACCTCTGGTGTATGATGCGCTGTGGCGCTATACACAGAGCTGGGTCTGG 905
Qy      281 AlaTyrrGluIleCysThrPheLeuLysAsnGlyAlaThrGlnGlyTrpAspAlaPro 300
Db      906 GCCTACTATGAGATTTGCACTTCTTGAGAAGTGGAGCCACTGAGGTCTGGGATGCTCC 965
Qy      301 GlnGluValProTyrrAlaTyrrGlnGlyAsnValTrpValGlyTyrrAspAsnIleLysSer 320
Db      966 CAGAAGTGCCTTATGCCATTAAGGCCAAGAGTGGCTTGGCTATGACAAATATCAGAGC 1025
Qy      321 PheAspIleLysAlaGlnTrpLeuLysHisAsnLysPheGlyGlyAlaMetValTrpAla 340
Db      1026 TTCAGGTGTTAAGGCTCAGTGGCTTAAGCAGAACAAATTTGGAGGTGCCATGATCTGGGCC 1085
Qy      341 IleAspLeuAspAspPheThrGlyThrPheCysAsnGlnGlyPheProLeuIleSer 360
Db      1086 ATTGACCTTGATGACTTCACTGGCTCTTCTGTGATCAGGAGAAATTTCTCTGACTTCT 1145
Qy      361 ThrLeuLysAlaLeuGlnGlyLeuGlnSerAlaSerCysThrAlaProAlaGlnProIle 380
Db      1146 ACTTTGAACAAGACCCCTGGCATTCATCCATGAAGTGTGACAGCTCCTGACGTGGCTTCC 1205
Qy      381 GluProIleThrAlaAlaProSerGlySerGlySerGlyAsnGlySerGlySerSerGly 400
Db      1206 GAGCCAGTACTACTCTCTCA-----GGAGTGGGAGTGGGGTGGAGGCTCGGA 1256
Qy      401 GlySerSerGlyGlySerGlyPheCysAlaValArgAlaAsnGlyLeuTyrrProValAla 420
Db      1257 GGAAGCTCTGGAGGCGAGTGGATTTCTGTGCGCAGAAAGCAGATGGCTCTACCTGTGGCA 1316
Qy      421 AsnAsnAtqAsnAlaPheTrpHisCysValAsnGlyValTyrrGlnGlnAsnCysGln 440
Db      1317 GATGCAAGAAATGCTTTTGGCAGTGCATCAATGAATCACAATACAGAGCATTTGTCAA 1376
Qy      441 AlaGlyLeuValPheAspThrSerCysAspCysCysAsnTrp 454
Db      1377 GCAGGCTGTTTGTGATACCACTGATGATTTGCTGCACTGG 1418

RESULT 8
AB051629 1529 bp mRNA linear MAM 07-NOV-2001
LOCUS   Bos taurus cbp b04 mRNA for chitin binding protein b04, complete
DEFINITION
ACCESSION AB051629
VERSION   AB051629.1 GI:16754894
KEYWORDS Bos taurus liver cDNA to mRNA.
SOURCE   Bos taurus
ORGANISM Bos taurus

REFERENCE
AUTHORS Suzuki,M., Morimatsu,M., Yamashita,T., Iwanaga,T. and Syuto,B.
TITLE A novel serum chitinase that is expressed in bovine liver
JOURNAL FEBS Lett. 506 (2), 127-130 (2001)
MEDLINE 21475601
REFERENCE 2 (bases 1 to 1529)
AUTHORS Suzuki,M., Morimatsu,M. and Syuto,B.
TITLE Direct Submission
JOURNAL Submitted (23-NOV-2000) Masako Suzuki, Iwate University, Department of Veterinary Medicine, Faculty of Agriculture; Ueda 3-18-8, Morioka, Iwate 020-8550, Japan (E-mail:u9998001@iwate-u.ac.jp, Tel:81-19-621-6212)
FEATURES Location/Qualifiers

```

```

source 1..1529
/organism="Bos taurus"
/db_xref="taxon:9913"
/tissue_type="liver"
gene 1..1529
/feature="cbp b04"
CDS 13..1431
/feature="cbp b04"
/codon_start=1
/product="chitin binding protein b04"
/protein_id="BAB71805.1"
/db_xref="GI:16754895"
/translation="MAKILPITGLAFLNLAQLGSAYLVYVPSNWAQVRPGLSGPKPD
NIDPLCLHLIYAFAGMNSITTTTENDVALYSSFNLDKKNSQLKILLATGWNFG
TAPFTAMVATPENRTFTSSVTKLHOYFGDLDFDWEVPGPRGSPSODKHLPTVLVQ
ETRAFEQEKOTNKPRLVTAAGAIGSLQAGVEIPLSQYLDPIHMTYDHPGSM
EYTGENSEPLKYPTDTSNTYLVVEYAMNWKNGAPAEKLIIGFPIYGNHFIIRDA
SNNGIYSGAGPAGPYTRRAGWAYEIECAFLKDGATGAEWDDSQNPYATYGTETWV
GYDNVNSPRIKAQWLKENFGAMVWALDDFTGTCFCNKKFPLINTLKDALGKSA
TCNASTQSEPNSSPNSGSGNSKSSSESGRGYAGKADGLYPVADNRNFAFNVCNGI
TYKQNLGLVFDTSCHCNNA"
polya_signal 1508..1513
/feature="cbp b04"
BASE COUNT 391 a 405 c 365 g 368 t
ORIGIN
Alignment Scores:
Pred. No.: 2,99e-148 Length: 1529
Score: 2039.00 Matches: 365
Percent Similarity: 86.53% Conservative: 29
Best Local Similarity: 80.22% Mismatches: 57
Query Match: 81.36% Indels: 4
DB: 4 Gaps: 1
US-10-004-219B-14 (1-455) x AB051629 (1-1529)
Qy 1 TyrGlnLeuThrCysTyrrPheThrAsnTrpAlaGlnTyrrArgProGlyLeuGlyArgPhe 20
Db 76 TACCAGCTGGTATGCTACTTCTTAACCTGGGCCAGTACCGTCCAGGCTGGAGCTTC 135
Qy 21 MetProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrrAlaPheAlaGlyArg 40
Db 136 AAGCCTGCAACATCGACCCCTCGCTCGCATCACTGATCTATGCTTCGCTGGAGT 195
Qy 41 GlnAsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrrGlnAlaPheAsn 60
Db 196 AGTAACAGTGAGATCACCACCAATGGAATGAATGATGGCTCTCTATAGTTCTTTCAT 255
Qy 61 GlyLeuLysAsnLysAsnSerGlnLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPhe 80
Db 256 GACCTGAAAAAAGAAAGACAGCCAGCTGAAAAATTCCTCTGGCCATTTGGAGCTGGAATTT 315
Qy 81 GlyThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThr 100
Db 316 GGAAGTGGCCCTTTTCACTGCTATGTTGGTCCACTCCGAGAACCGAAGACTTTTCTTCT 375
Qy 101 SerValIleLysPheLeuArgGlnTyrrGluPheAspGlyLeuAspPheAspTrpGluTyrr 120
Db 376 TCAGTCATCAAAATTCCTACACCATGATGGATTTGACGGCTGGATTTGATTTGGAGATAC 435
Qy 121 ProGlySerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValGlnGlu 140
Db 436 CTTGCTTTTGGTGAAGCCCTTCTCAGGACCAAGCATCTTCTTCACTGTCTTAGTGAGGAA 495
Qy 141 MetArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThr 160
Db 496 ACACGTGAAGCTTTTGAACAGAGAGCAAGCAACCAACAGCCAGGCTGCTGTGTCC 555
Qy 161 AlaAlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrrGluIleProGlnLeuSer 180
Db 556 GCTGAGTAGTGTCTGGCATCTCTAATATCCAGGCTGGCTATGAGATTTCCAGCTATCC 615
Qy 181 GlnTyrrLeuAspTyrrIleHisValMetThrTyrrAspLeuHisGlySerTrpGluGlyTyrr 200

```

```
Db 616 CAGTACCTGGACTTCCATGTCATGCTATGACTTCCATGGCTCTCTGGGAAGGCTAC 675
Qy 201 ThrGlyGluAsnSerProLeuTyrLysTyrProThrAspThrGlySerAsnAlaTyrLeu 220
Db 676 ACAGGAGAACACGCCCTCTCTACAATACCCAACTGACACTGGCAGCAATACTTACCTTC 735
Qy 221 AsnValAspTyrValMetAsnTyrTrpLysAspAsnGlyValAlaProAlaGluLysLeuIle 240
Db 736 AATGTGGAATACGCCATGAACCTACTGGAAGAAAATGGGCCCGCTGAGAAGCTCATC 795
Qy 241 ValGlyPheProThrTyrGlyHisAsnPhelIleLeuSerAsnProSerAsnThrGlyIle 260
Db 796 ATTGGATTCCCTGCTACGACACAATTTTCATCCTAAGAGAGCTTCCCAACAATGGTATC 855
Qy 261 GlyAlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyIleTTP 280
Db 856 GGTGCCCCCATTCTGGAGTGTGCTCTCGCGGCCCTACACGAGGAGGCTGGGTCTGG 915
Qy 281 AlaTyrTyrGluIleCysThrPheLeuLysAsnGlyAlaThrGlnGlyTrpAspAlaPro 300
Db 916 GCCTACTATGAGATCTGTGCTTCTTGAAAGATGGAGCCACTGAGGCATGGGATGACTCT 975
Qy 301 GlnGluValProTyrAlaTyrGlnGlyAsnValTrpValGlyTyrAspAsnIleLysSer 320
Db 976 CAGATGTGCTCTATGCTACAGAGGCACTGAGTGGGTGGTTATGACAATGTTAACAGC 1035
Qy 321 PheAspIleValAlaGlnTrpLeuLysHisAsnLysPheGlyValMetValTrpAla 340
Db 1036 TTCAGAAATCAAGGCTCAGTGGCTTGAAGGAGAACTTTGGAGGTGCCATGGTCTGGGCC 1095
Qy 341 IleAspLeuAspAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSer 360
Db 1096 ATTGACTTGGATGTTTCACAGGCATCTTTCGCAACAGGCGCAAGTTCCTCCCTGATAAT 1155
Qy 361 ThrLeuLysLysAlaLeuGlyLeuGlnSerAlaSerCysThrAlaProAlaGlnProIle 380
Db 1156 ACCTTGAAGGATCTCTGCGCTGGAAGAGTGCAACCTGTAATGCCCTCCAGCAGTCCAGT 1215
Qy 381 GluProIleThrAlaAlaProSerGlySerGlyAsnGlySerGlySerSerSerGly 400
Db 1216 GAGCCCC-----AACACAGCTCTCTGGAAACGAAGTGGAGTGGAAATAGAGC 1263
Qy 401 GlySerSerGlyGlySerGlyPheCysAlaValAlaGlnGlyLeuTyrProValAla 420
Db 1264 AGCAGCTCCGAGGTAGAGATCTGTGCTGGCAAGCTGATGCTCTTACCCAGTGGCG 1323
Qy 421 AsnAsnArgAsnAlaPheThrHisCysValAsnGlyValThrTyrGlnGlnAsnCysGln 440
Db 1324 GATAACAGAAATGCCCTTCTGGAATCGTAAATGGAATCACAATACAAACAGAAATTCCTG 1383
Qy 441 AlaGlyLeuValPheAspThrSerCysAspCysAsnTrpAla 455
Db 1384 ACTGGGCTTGTCTTTGACACCACTGTCTACTGTCTGCAACTGGGCA 1428

RESULT 9
LOCUS AB025009
DEFINITION Homo sapiens TSA1902-S mRNA for novel member of chitinase family,
complete cds.
ACCESSION AB025009
VERSION AB025009.1 GI:6467178
KEYWORDS novel member of chitinase family; TSA1902-S.
SOURCE Homo sapiens cDNA to mRNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (sites)
REFERENCE Saito, A., Ozaki, K., Fujiwara, T., Nakamura, Y. and Tanigami, A.
AUTHORS Isolation and mapping of a human lung-specific gene, TSA1902,
TITLE encoding a novel chitinase family member
JOURNAL Gene 239 (2), 325-331 (1999)
MEDLINE 20018184
```

```
2 (bases 1 to 1188)
REFERENCE Saito, A., Ozaki, K., Fujiwara, T., Takahashi, E. and Tanigami, A.
AUTHORS Direct Submission
TITLE JOURNAL
SUBMITTED (16-MAR-1999) Akihiko Saito, Otsuka Pharmaceutical Co.,
Ltd., Otsuka GEN Research Institute; 463-10 Kagasuno Kawauchi-cho,
Tokushima, Tokushima 771-0192, Japan (E-mail:saito@otsuka.gr.jp,
Tel:81-88-665-2888, Fax:81-88-637-1035)
FEATURES
source
1..1188
/organism="Homo sapiens"
/db_xref="taxon:9606"
gene
1..1188
/gene="TSA1902-S"
CDS
150..1097
/gene="TSA1902-S"
/codon_start=1
/product="novel member of chitinase family"
/protein_id="BAA86981.1"
/db_xref="GI:6467179"
/translation="MREAFQEAKQINPRLMTAAVAAGISNIOGVEIPLSOYLD
YIHTYDILHSGWEGTGTGNSPLKYPTDTSNAYLVNVDYVYVWYKNGAPAEKLI
PFTYGNFILSPNTGIGTSGAGPAGPAGKSGIWAYVEICTFLKNGATQGWDA
QEVPAVQGVNVDYVKSFDIKQWLKHNKFGAMVWAILDDFTCTFCNQKPEL
ISTLKALGLOSASCTAPQPIETPITAAPSGSGSGSSGSGSGSGGFCNAGRAN
L
YEVANNRNFHVCVNGVTYQNCQALVFDTSDDCNWA"
BASE COUNT 280 a 337 c 309 g 262 t
ORIGIN
Alignment Scores: 1.61e-127 Length: 1188
Score: 1768.50 Matches: 331
Percent Similarity: 85.13% Conservative: 1
Best Local Similarity: 84.87% Mismatches: 3
Query Match: 70.57% Indels: 56
DB: 9 Gaps: 1
US-10-004-219B-14 (1-455) x AB025009 (1-1188)
Qy 66 AsnSerGlnLeuLysThrLeuLeuAlaIleGlyGlyTyrAsnPheGlyThrAlaProPhe 85
Db 91 AACAGCCAGCTGAANAACCTCTCTGGCCATTGGAGCTGGAACCTTCAGGACTGC-CCC--- 146
Qy 86 ThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThrSerValIleLysPhe 105
Db 146 ----- 146
Qy 106 LeuArgGlnTyrGluPheAspGlyLeuAspPheAspTyrGluTyrProGlySerArgGly 125
Db 146 ----- 146
Qy 126 SerProGlnAspLysHisLeuPheThrValLeuValGlnGluMetArgGluAlaPhe 145
Db 147 -----GAAATGCGTGAAGCTTTT 164
Qy 146 GluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThrAlaAlaValAlaIle 165
Db 165 GAGCAGAGCGCCCAAGCAGATCAACAAGCCAGGCTGATGCTGCTGCTGCTGCTGCTGCT 224
Qy 166 GlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSerGlnTyrLeuAspTyr 185
Db 225 GGCATCTCCAAATATCCAGTCTGGCTATGAGATCCCCCAACTGCACAGTACTGACTGAC 284
Qy 186 IleHisValMetThrTyrAspLeuHisGlySerTyrGluGlyTyrThrGlyGluAsnSer 205
Db 285 ATCCATGTCATGACCTACGACCTCCATGGCTCTCTGGAGGCTACACTGGAGAGACAGC 344
Qy 206 ProLeuTyrLysTyrProThrAspThrGlySerAsnAlaTyrLeuAsnValAspTyrVal 225
Db 345 CCCCTCTCAAAATACCGCACTGACACCGCGCAGCAACGCCCTACTCAATGTGATATTGTC 404
Qy 226 MetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeuIleValGlyPheProThr 245
Db 405 ATGAACTACTGGNAGGACCAATGGAGCACAGCTGAGAGCTCATGTTGGATTCCCTACC 464
```

246 TyrGlyHisAsnPhelLeuSerAsnProSerAsnThrGlyValAlaProThrSer 265
 465 TATGACACAACTTCTCTGAGCAACCCCTCCACACTGGATTGGTGGCCCACTCT 524
 266 GlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyLeuTyrAlaTyrGluLeu 285
 525 GGTGCTGCTCTCTCTGGGCTTATGCAAGAGTCTGGGATCTGGGCTTACTACGAGATC 584
 286 CysThrPheLeuLysAsnGlyAlaThrGlnGlyTyrAspAlaProGlnValProTyr 305
 585 TGTACTCTCTGAAATAATGGACCACTAGGAGTGGATGGCCCTCAGGAATGGCTTAT 644
 306 AlaTyrGlnGlyAsnValTyrValGlyTyrAspAsnLeuLysSerPheAspLeuLysAla 325
 645 GCTATCAGGGCAATGCTGGGTGGCTATGACACGTCACAGCTTCGATTAAGGCT 704
 326 GlnTyrPheLeuLysHisAsnLysPheGlyGlyAlaMetValTyrAlaLeuAspLeuAsp 345
 705 CAATGGCTTAAGCACAACTTGGAGCGGCTGCTGGGCCATTGATCTGGATGAC 764
 346 PheThrGlyThrPheCysAsnGlnGlyLysPheProLeuLeuSerThrLeuLysLysAla 365
 765 TTCCTGAGGCACTTCTTGCAACAGGCGCAAGTTCCCTTAATCTCCACCTGAGGAGGCC 824
 366 LeuGlyLeuGlnSerAlaSerCysThrAlaProAlaGlnProIleGluProIleThrAla 385
 825 CTTGGCTGCAGTGCAAGTTGCAGGCTCCAGCTCAGCCATTGAGCCATTAAGTCT 884
 386 AlaProSerGlySerGlyAsnGlySerGlySerSerSerGlyGlySerGlyGly 405
 885 GCTCCAGTGGCAGCGGAAACGGAGCGGAGTAGCAGCTCTGGAGGCACTCGGAGGC 944
 406 SerGlyPheCysAlaValAlaGlnGlyLeuTyrProValAlaAsnArgGlnAla 425
 945 AGTGGATCTGTGCTGGCAGACCAACGCGCTCTACCCGTCGCAATTAACAGAAATGCC 1004
 426 PheTyrPheCysValAsnGlyValThrTyrGlnGlnAsnCysGlnAlaGlyLeuValPhe 445
 1005 TTCCTGGCACTGGTGAATGAGTCACTACGACAGCAAGTGCAGCGCGGCTGTCTTC 1064
 446 AspThrSerCysAspCysCysAsnTyrAla 455
 1065 GACACAGCTGTGATGCTGCAACTGGGCA 1094

AK098814 4250 bp mRNA linear PRI 15-JUL-2002
 Homo sapiens cDNA FLJ25948 fis, clone STM04207, highly similar to
 Homo sapiens acidic mammalian chitinase precursor.
 AK098814
 AK098814.1 GI:21758941
 Cligo capping; fis (full insert sequence).
 Homo sapiens stomach mucosa cDNA to mRNA, clone_lib:STM
 clone:STM04207.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

TITLE
 NEDO human cDNA sequencing project
 JOURNAL
 Unpublished
 REFERENCE
 2 (bases 1 to 4250)
 Sugano, S. and Suzuki, Y.
 Direct Submission
 TITLE
 Submitted (08-JUL-2002) Sumio Sugano, Institute of Medical Science,
 University of Tokyo, Laboratory of Genome Structure, Human Genome

Center; Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
 (E-mail:cdna@ims.u-tokyo.ac.jp, Tel.81-3-5449-5286,
 Fax:81-3-5449-5416)
 NEDO human cDNA sequencing project supported by Ministry of
 Economy, Trade and Industry of Japan; cDNA full insert sequencing:
 Research Association for Biotechnology (RAB); cDNA library
 construction and 5'-end one pass sequencing: Institute of Medical
 Science, University of Tokyo, Laboratory of Genome Structure, Human
 Genome Center; 3'-end one pass sequencing: RAB; clone selection for
 full insert sequencing: RAB and Helix Research Institute.

FEATURES

source
 1..4250
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="STM04207"
 /tissue type="stomach mucosa"
 /clone_lib="STM"
 /note="cloning vector: pME18SFL3"
 BASE COUNT 1262 a 868 c 966 g 1154 t
 ORIGIN

Alignment Scores:

Pred. No.: 8,69e-125 Length: 4250
 Score: 1742.00 Matches: 318
 Percent Similarity: 98.47% Conservative: 3
 Best Local Similarity: 97.55% Mismatches: 5
 Query Match: 69.51% Indels: 0
 DB: 9 Gaps: 0

US-10-004-219B-14 (1-455) x AK098814 (1-4250)

Qy 130 AspLysHisLeuPheThrValLeuValGlnGluMetArgGluAlaPheGluGlnGluAla 149
 Db 3174 GATGTCCTATTATGCTTATTCTCTAGGAAATGCGTGAAGCTTTTGAGCAGGAGGCC 3233
 Qy 150 LysGlnIleAsnLysProArgLeuMetValThrAlaAlaValAlaAlaGlyIleSerAsn 169
 Db 3234 AAGCAGATCAACAAGCCAGGCTGATGGTCACTGCTGCAGTAGCTGCTGGCATCTCCAAT 3293
 Qy 170 IleGlnSerGlyTyrGluIleProGlnLeuSerGlnTyrLeuAspTyrIleHisValMet 189
 Db 3294 ATCCAGTCTGGCTATGAGATCCCACTGTACAGTAGTACCTGGACTACATCATGTGCTG 3353
 Qy 190 ThrTyrAspLeuHisGlySerTyrGluGlyTyrThrGlyGluAsnSerProLeuTyrLys 209
 Db 3354 ACCTAGACCTCCATGGCTCTCTGGAGGGCTACACTGGAGAGAACAGCCCTCTACAA 3413
 Qy 210 TyrProThrAspThrGlySerAsnAlaTyrLeuAsnValAspTyrValMetAsnTyrTyr 229
 Db 3414 TACCCGACTGACACGGCAGCAACCGCTACCTCAATGTGGATTATGTCATGAACACTGCTG 3473
 Qy 230 LysAspAsnGlyAlaProAlaGlyLysLeuIleValGlyPheProThrTyrGlyHisAsn 249
 Db 3474 AAGGCAATGGAGCACCAGCTGAGAGGCTCAATGTGGATTCCCTACCTATGGACACAAC 3533
 Qy 250 PheIleLeuSerAsnProSerAsnThrGlyIleGlyAlaProThrSerGlyAlaGlyPro 269
 Db 3534 TTCATCTCTGAGCAACCCCTCCAACTGGAATGGTGGCCCACTCTGCTGCTGCTCT 3593
 Qy 270 AlaGlyProTyrAlaLysGluSerGlyIleTyrAlaTyrTyrGluIleCysThrPheLeu 289
 Db 3594 GCTGGGCCCTATGCCAAGGAGTCTGGGCTTACTAGGAGATCTGTACCTTCTCTG 3653
 Qy 290 LysAsnGlyAlaThrGlnGlyTyrAspAlaProGlnGluValProTyrAlaTyrGlnGly 309
 Db 3654 AAAAAATGGAGCCACTCAGGAGTGGATGCCCTCAGGAAGTGCCTTATGCCCTATCAGGGC 3713
 Qy 310 AsnValTyrValGlyTyrAspAsnIleLysSerPheAspIleLysAlaGlnTyrLeuLys 329
 Db 3714 AATGTGGTGGTGGCTATGACAACTCAAGACTTCGATATTAAAGCTCAATGGCTTAAG 3773
 Qy 330 HisAsnLysPheGlyGlyAlaMetValTyrAlaIleAspLeuAspPheThrGlyThr 349


```
Db 3774 CACAACAAATTTGGAGCGCCATGTTCTGGGCCATTGATCTGGATGACTTCACTGGCACT 3833
Qy 350 PheCysAsnGlnGlySerPheProLeuIleSerThrLeuIleValAlaLeuGlyLeuGln 369
Db 3834 TTCGTCAACAGGCGCAAGTTTCCCTTAATCTCCACCCTGAAGAAGCCCTCGCCCTGGAG 3893
Qy 370 SerAlaSerCysThrAlaProAlaGlnProIleGluProIleThrAlaAlaProSerGly 389
Db 3894 AGTGCAGTTGACACGGCTCCAGCTCAGCCATTGAGCCATAACTGCTCTCCAGTGGC 3953
Qy 390 SerGlyAsnGlySerGlySerSerSerGlySerGlySerGlySerGlySerGlyPheCys 409
Db 3954 AGCGGAAACGGAGCGGAGTAGCAGCTCTGGAGGAGCTCGGGAGCGAGTGGATTCTGT 4013
Qy 410 AlaValArgAlaAsnGlyLeuTyrProValAlaAsnAsnArgAsnAlaPheTrpHisCys 429
Db 4014 GCTGTGAGAGCAACCGGCTCTACCCCGTGGCAATTAACAGAAATGCCCTTCTGGCACTGC 4073
Qy 430 ValAsnGlyValThrTyrGlnGlnAsnCysGlnAlaGlyLeuValPheAspThrSerCys 449
Db 4074 GTGAATGGAGTCACGTACACGAGCAACTGCCAGGCGGGCTTGTCTTCGACACCAGCTGT 4133
Qy 450 AspCysCysAsnTrpAla 455
Db 4134 GATTGCTCAACTGGGCA 4151.

RESULT 11
M94584
LOCUS M94584 1525 bp mRNA linear ROD 15-JUN-2001
DEFINITION Mus musculus secretory protein precursor (Yml) mRNA, complete cds.
ACCESSION M94584
VERSION M94584.2 GI:11140876
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1525)
AUTHORS Chang, N.-C.A., Liu, C.-H. and Chang, A.C.
TITLE Molecular characterization of a secretory protein (YM-1)
JOURNAL transiently expressed by activated murine peritoneal macrophages
REFERENCE 2 (bases 1 to 1525)
AUTHORS Chang, N.-C., Hung, S.I., Hwa, K.Y., Kato, I., Chen, J.E., Liu, C.H. and Chang, A.C.
TITLE A macrophage protein, Yml, transiently expressed during
JOURNAL inflammation is a novel mammalian lectin
MEDLINE J. Biol. Chem. 276 (20), 17497-17506 (2001)
PUBMED 21264517
REFERENCE 3 (bases 1 to 1525)
AUTHORS Chang, N.-C.A., Liu, C.-H. and Chang, A.C.
TITLE Direct Submission
JOURNAL Submitted (27-APR-1993) Microbiology and Immunology, National
Yang-Ming University, 155, Sect. II, Li-Noon St, Taipei, Taiwan
112, R.O.C.
REFERENCE 4 (bases 1 to 1525)
AUTHORS Chang, N.-C.A., Liu, C.-H. and Chang, A.C.
TITLE Direct Submission
JOURNAL Submitted (13-NOV-2000) Microbiology and Immunology, National
Yang-Ming University, 155, Sect. II, Li-Noon St, Taipei, Taiwan
112, R.O.C.
REMARK Sequence update by submitter
COMMENT On Nov 13, 2000 this sequence version replaced gi:202441.
FEATURES
Location/Qualifiers
1..1525
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone_lib="PEC in lambda gt11 and ZAPII"
/dev_stage="adult"
1..1525
/gene="Yml"
13..1209
CDS
```

```
/gene="Yml"
/codon_start=1
/product="secretory protein precursor"
/protein_id="AAB62394.2"
/db_xref="GI:11140877"
/translation="MAKILVTGLAILNLVQLSSYQLMCMYTTWAKDRPIEGSPKPG
NIDPLCTHLIYAFAGMONEITTHEQDLRDEALNGDKNTKLTLLAIGGWKFG
PAPFSAMYSTPQNRQIFQSVIRFLROYNFDGLNLDWQYPSRSGSPDKHLFSLVK
EMRKAPESEVEKDI PRLLTSTGAGI IDVIKSGYKI PELQSOLDYI OVMYTDLHDPK
DGYTGNSPLYKSPVDIGKSADLANVDSIISYMKDHGAASEKLVGFPAYGHTFILLSDP
SKTGICARTISTGPPCKYTDSSGLLAYEVECTFLNEGATEWDAPOEVPYAYQGNWV
GYDNVRSFKLAQWLKNNLGGAVVWPLDMDDFSGFCHQRHFPLTSLKGLDNIHSA
SCKGPY"
sig_peptide 13..75
mat_peptide 76..1206
polyA_signal /product="secretory protein"
polyA_site /gene="Yml"
BASE COUNT 434 a 343 c 335 g 413 t
ORIGIN
Alignment Scores:
Pred. No.: 4..1e-100 Length: 1525
Score: 1415.00 Matches: 256
Percent Similarity: 80.85% Conservative: 48
Best Local Similarity: 68.09% Mismatches: 72
Query Match: 56.46% Indels: 0
DB: 10 Gaps: 0
US-10-004-219B-14 (1-455) x M94584 (1-1525)
Qy 1 TyrGlnLeuThrCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPhe 20
Db 76 TACCAGCTGATGTGCTACTATACCACTGGGTAGGACAGGCCCAATAGAGGGAGTTTC 135
Qy 21 MetProAspAsnIleAspProCysIleuCysThrHisLeuIleTyrAlaPheAlaGlyArg 40
Db 136 AAACCTGGTAATATTGACCCCTGCTGTACTACCTGATCTATGCTCTTCTGGAATG 195
Qy 41 GlnAsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrGlnAlaPheAsn 60
Db 196 CAGATAATGAGATCATTACACACATGACGAGCAAGACTTGGCGTACTATGAGCAATTGAAT 255
Qy 61 GlyLeuLysAsnLysAsnSerGlnLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPhe 80
Db 256 GGTCTGAAAGACAGACACACTGAGCTGAGCTGCTCTGCGCCATTGGAGGATGGAAGTTT 315
Qy 81 GlyThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThr 100
Db 316 GGACCTGCCCCGTTTCAGTGCCATGTGCTCTACTCTCAGAACCGTCAGATATTCATTTCAG 375
Qy 101 SerValIleLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyr 120
Db 376 TCAGTTATCAGATTCCTTCGTCATATACTTTGATGGCTCAACCTGAGCTGGCAGTAC 435
Qy 121 ProGlySerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValGlnGlu 140
Db 436 CCTGGGTCCTCGAGGAAGCCCTCTTAAGGACAAACATCTCTTCAGTGTTCTGTGTGAAGAA 495
Qy 141 MetArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThr 160
Db 496 ATGCGTAAGCTTTTGAGGAAGAAATCTGTGGGAAGAACATTCGAGGCTGCTCTACT 555
Qy 161 AlaAlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSer 180
Db 556 TCCACAGGAGCAGGAATCATTGACGCTAATCAAGTCTGGGTACAGATCCCTCAACTGCT 615
Qy 181 GlnTyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr 200
Db 181
```



```
Db 616 CAGTCTCTTGACTATATTACAGTCATGACATATGATCTCCATGATCCTTAAGGATGGCTAC 675
Qy 201 ThrGlyGluAsnSerProLeuTyrIleYsTyrProThrAspThrGlySerAsnAlaTyrLeu 220
Db 676 ACTGGAGAAAATAGTCCCTCTATAAATCTCATATGACATTTGGAAGAGTCTGATCTC 735
Qy 221 AsnValAspTyrValMetAsnTyrTrpIysAspAsnGlyAlaProAlaGluIysLeu 240
Db 736 AATGTGGATTCAATCAATTTCTACTGGAAGGACCATGGAGCAGCTTCTGAGAAGCTCAT 795
Qy 241 ValGlyPheProThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIle 260
Db 796 GTGGGATTTCAGCATATAGGATACCTTTATCTGAGTGACCCCTTCTAAGACTGGAAT 855
Qy 261 GlyAlaPheThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyIleTrp 280
Db 856 GTGGCCCTACATAGTAGTACTGGCCACAGGAAGTACACAGATGAATCAGAGATCTCTG 915
Qy 281 AlaTyrTyrGluIleCysThrPheLeuIysAsnGlyAlaThrGlnGlyTrpAspAlaPro 300
Db 916 GCTTACTATGAGGTTGTACATTTCTGAATGAAGGAGCCACTGAGGTCTGGGATGCCCC 975
Qy 301 GlnGluValProTyrAlaTyrGlnGlyAsnValTrpValGlyTyrAspAsnIleYsSer 320
Db 976 CAGGAAGTACCTATGCTTATCAGGGTAATGAGTGGTGTGTTATGACAAATCTCAGGAGC 1035
Qy 321 PheAspIleLysAlaGlnTrpLeuIysHisAsnLysPheGlyGlyAlaMetValTrpAla 340
Db 1036 TTCAGTGTGAAGGCTCAGTGGCTCAAGGACAATTTAGGAGGTGCCGTCTGGCC 1095
Qy 341 IleAspLeuAspAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSer 360
Db 1096 CTGGACATGATGATCTCAGTGGTCTTCTCTGTCACCAGAGACATTTCCCTCTGACATCT 1155
Qy 361 ThrLeuLysLysAlaLeuGlyLeuGlnSerAlaSerCysThrAlaPro 376
Db 1156 ACTTTAAAGGAGATCTCAATATACACAGTGAAGTTGCAAGGGCCCT 1203

RESULT 12
Db 87757 1506 bp mRNA linear ROD 12-JAN-2000
DEFINITION Mus musculus mRNA for ECF-L precursor, complete cds.
ACCESSION D87757
VERSION D87757.2 GI:6015435
KEYWORDS ECF-L precursor.
SOURCE Mus musculus bone marrow cDNA to mRNA.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Ohashi, M., Arita, H. and Hwai, N.
1 (bases 1 to 1506)
Identification of a novel eosinophil chemotactic cytokine (ECF-L)
as a chitinase family protein
J. Biol. Chem. 275 (2), 1279-1286 (2000)
REFERENCE 2 (bases 1 to 1506)
Ohashi, M.
Direct Submission
Submitted (06-SEP-1996) Makoto Ohashi, The University of
Tokushima, Faculty of Integrated Arts and Sciences; 1-1
Minami-Johsanjima, Tokushima, Tokushima 770, Japan
(E-mail: ohashi@ias.tokushima-u.ac.jp, Tel: 0886-56-7261,
Fax: 0886-56-7298)
On Oct 7, 1999 this sequence version replaced gi:1545818.
Sequence location (05-Oct-1999).
FEATURES
source
1..1506
/organism="Mus musculus"
/db_xref="taxon:10090"
/tissue_type="bone marrow"
1..1506
/gene="mECF-L"
1..1197
CDS
```

```
/gene="mECF-L"
/codon_start=1
/product="ECF-L precursor"
/protein_id="BAAL3458.2"
/db_xref="GI:6015435"
/training="MAKULITVGLAILNVLGSSYQLMCMYTSWAKORPIEGSEKPG
NIDPCLTHLIYAFAGMONEITTHODLRYEALNGKDKNTKLTLLAIGWKFG
PASTSAVSTPQNRQIFQSVIRFLRQYDFDLNLMQYPGSRGSPDKHLFSLVK
EMKAFBESVEKDIPLRLITSTGAGIIVDKSGYKIPELSQSLDYIOVMTVDLHDPK
DYGATGSPKSPVDIGKSLADLVDSIISYWKDHGAASEKLIIVGFPAVHTFILSDP
SKTGIGAPTISTGPGKYTDESGLLAYVEVCTFLNEGATEVMDAPOEVPYAYQGNW
GYNVRFSFKLAKQMLKONNLGAVVWPLMDMDFSGSFCHQRHFLPLTSLKGLDLNHS
SCKGPY"
BASE COUNT 423 a 338 c 333 g 412 t
ORIGIN
Alignment Scores:
Pred. No.: 1506
Score: 1407.00 Matches: 255
Percent Similarity: 80.59% Conservative: 48
Best Local Similarity: 67.82% Mismatches: 73
Query Match: 56.15% Indels: 0
DB: 10 Gaps: 0
US-10-004-219B-14 (1-455) x D87757 (1-1506)
Qy 1 TyrGlnLeuThrCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPhe 20
Db 64 TACCAGCTGATGCTGCTACTATACAGTGGCTAGGACAGGCCAATAGAGGGAGTTTC 123
Qy 21 MetProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArg 40
Db 124 AAACCTGCTAATATTGACCCCTGCTGTACTACCTGATCTATGCTTTGCTGGAATG 183
Qy 41 GlnAsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrGlnAlaPheAsn 60
Db 184 CAGATAATAGATCACTTACACACATGACGACGACTTGCCTGACTATGACATTTGAAT 243
Qy 61 GlyLeuLysAsnLysAsnSerGlnLeuLysThrLeuAlaIleGlyGlyTrpAsnPhe 80
Db 244 GGTCTGAAGACAGACACATGAGCTAAACTCTCTCGCCATTTGGAGATGGAATTT 303
Qy 81 GlyThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThr 100
Db 304 GGACCTGCTCGTTCAGTGCCATGCTCTACTCTCTCAGAACCGTCAGATATTCATTAC 363
Qy 101 SerValIleLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyr 120
Db 364 TCAGTTATCAGATTCCTTCGTCATATAAATTTGATGGCTCAACCTGGAGTAC 423
Qy 121 ProGlySerArgGlySerProGlnAspLysHisLeuPheThrValLeuValGlnGlu 140
Db 424 CTGGGCTCTGAGGAGGCCCTCTTAAGGACAAACATCTCTTCAGTGTCTGTTGAGGAA 483
Qy 141 MetArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThr 160
Db 484 ATGCTGAAGCTTTTGAAGAGAAATCTGTGGAGAAAGACATTTCCAAGGCTGCTACT 543
Qy 161 AlaAlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSer 180
Db 544 TCCACAGGAGCAGGAATCAATTCAGCGTAATCAAGTCTGGGTCAAGATCCCTGAACT 603
Qy 181 GlnTyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr 200
Db 604 CAGTCTCTTGACTATATTACAGTCAATGATATGATCTCCATGATCCTTAAGATGGCTAC 663
Qy 201 ThrGlyGluAsnSerProLeuTyrIleYsTyrProThrAspThrGlySerAsnAlaTyrLeu 220
Db 664 ACTGGAGAAAATAGTCCCTCTATAAATCTCATATGACATTTGGAAGAGTCTGATCTC 723
Qy 221 AsnValAspTyrValMetAsnTyrTrpIysAspAsnGlyAlaProAlaGluIysLeu 240
Db 724 AATGTGGATTCAATCAATTTCTACTGGAAGGACCATGGAGCAGCTTCTCAGAAAGCTCAT 783
```

```
QY 241 ValGlyPheProThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIle 260
Db 784 GTGGGATTTCCAGCATATGGCCATACCTTTATCTCTGAGTGACCTTCTTAAGACTGGAAT 843
QY 261 GlyAlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyIleTrp 280
Db 844 GGTGCCCTTACAAATTAGTAGTCTGCCCCACCAGGAAAGTACACATGAATCAGGACTCCTG 903
QY 281 AlaTyrTyrGluIleCysThrPheLeuLysAsnGlyAlaThrGlnGlyTrpAspAlaPro 300
Db 904 GCTTACTATGAGGTTTGTACATTTCTGAATGAAGGAGCCACTGAGTCTGGGATGCCCC 963
QY 301 GlnGluValProTyrAlaTyrGlnGlyAsnValTrpValGlyTyrAspAsnIleLysSer 320
Db 964 CAGGAAGTACCCTATGCTATCAGGTAATGAGTGGTGGTTGTTATGACAATGTCAGGAGC 1023
QY 321 PheAspIleLysAlaGlnTrpLeuLysHisAsnLysPheGlyAlaMetValTrpAla 340
Db 1024 TTCAAGTTGAAGGCTCAGTGGCTCAAGGACAAACAATTTAGGAGGTGCCGTGGTCTGCC 1083
QY 341 IleAspLeuAspAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSer 360
Db 1084 CTGACATGATGATCCTTCAAGTGGTCTTCTGTCACAGAGACATTTCCCTCTGACATCT 1143
QY 361 ThrLeuLysLysAlaLeuGlyLeuGlnSerAlaSerCysThrAlaPro 376
Db 1144 ACTTTAAGGAGATCTCAATATACACAGTGCAGTTCGCAAGTTCGCAAGGCCCT 1191
RESULT 13
LOCUS AY049765
DEFINITION Mus musculus secreted protein precursor Ym2 mRNA, complete cds.
ACCESSION AY049765
VERSION AY049765.2 GI:22123906
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1209)
AUTHORS Webb,D.C., McKenzie,A.N. and Foster,P.S.
TITLE Expression of the Ym2 lectin-binding protein is dependent on
interleukin (IL)-4 and IL-13 signal transduction: identification of
a novel allergy-associated protein
J. Biol. Chem. 276 (45), 41969-41976 (2001)
JOURNAL 21551268
MEDLINE 11553626
PUBMED 2 (bases 1 to 1209)
AUTHORS Webb,D.C. and Foster,P.S.
TITLE Direct Submission
JOURNAL Submitted (02-AUG-2001) Biochemistry and Molecular Biology, John
Curtin School of Medical Research, Australian National University,
Off Mills Rd, Acton, Canberra, ACT 2601, Australia
REFERENCE 3 (bases 1 to 1209)
AUTHORS Webb,D.C. and Foster,P.S.
TITLE Direct Submission
JOURNAL Submitted (06-AUG-2002) Biochemistry and Molecular Biology, John
Curtin School of Medical Research, Australian National University,
Off Mills Rd, Acton, Canberra, ACT 2601, Australia
REMARK Sequence update by submitter
COMMENT On Aug 6, 2002 this sequence version replaced gi:15705156.
FEATURES
source 1..1209
/organism="Mus musculus"
/strain="BALB/c"
/db_xref="taxon:10090"
/notes="isolated from the lung lavage from allergic mice"
1..1209
/notes="expression regulated by IL-4 and IL-13 signaling
via the IL-4R alpha subunit"
/codon_start=1
/product="secreted protein precursor Ym2"
```

```
/protein_id="AAL03953.2"
/db_xref="GI:22123907"
/translation="MAKILVLTGLAILNLVOLGSSYQLMCMYTSWAKDRPTGSPKPG
NIDPCLCTHLIVAFAGMKNNEITLSEODLYEALNGKDRNLTGLKLLAIGGWKFG
PADPSSMYSVTPQNRQTEIKSVIRFLRYNFDGLNDWOYPGSRGSPDKDHLFSVLVQ
EMRKAFESSTLNHILPRLLSTGAGFDIVIKSGYKIFELSQSLDIYIQWTFYDLHDPK
NGYTGENSEPLYSKYDYGKADNLNDVSIITTKDHGAASEKLIVGFPGAYHTFLLSDP
SKNGIGDPTVSAGPGKYTNFQGLAYFEICTFLNEGATEIFDQFVFPYAYLGNBW
GYDNVRSFKLKAQMLKDNLNGAVVWPLDMDDFSGSFCHQGRFPLTTTLTKRDLNVHSA
SCKASRYGEL"
sig_peptide 1..63
BASE COUNT 326 a 273 c 282 g 328 t
ORIGIN
Alignment Scores:
Pred. No.: 7,56e-98 Length: 1209
Score: 1384.00 Matches: 250
Percent Similarity: 80.53% Conservative: 52
Best Local Similarity: 66.67% Mismatches: 73
Query Match: 55.23% Indels: 0
DB: 10 Gaps: 0
US-10-004-219B-14 (1-455) x AY049765 (1-1209)
QY 1 TyrGlnLeuThrCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPhe 20
Db 64 TACCAGCTGATGTGCTACTATACAGCTGGGCTAAGGACAGCCCAACAGAGGGAGTTTC 123
QY 21 MetProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArg 40
Db 124 AAACCTGGTAATATTGACCCCTGCTGTGCTACCTGATCTATGCCCTTTCTCTGGGATG 183
QY 41 GlnAsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrGlnAlaPheAsn 60
Db 184 AAGATATATGATGATCATTACTTAAGTGAGCAAGACTTGGTGATGATGAACATTAAT 243
QY 61 GlyLeuLysAsnLysAsnSerGlnLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPhe 80
Db 244 GGTCTGAAAGACAGGACACACTGAGCTAAACCTCTCTGGCCATTGGAGGATGGAAGTTT 303
QY 81 GlyThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThr 100
Db 304 GGACTGCCCTCCCTCAGTTCCATGCTCTACTCTCAGAACCCGTCAGACATTCATTAAAG 363
QY 101 SerValIleLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyr 120
Db 364 TCAGTTATCAGATTCCTTCGTCAATATAACTTTGATGGCTCAACCTGAGCTGGCAGTAC 423
QY 121 ProGlySerArgGlySerProGlnAspLysHisLeuPheThrValLeuValGlnGlu 140
Db 424 CCTGGGTCTCGAGGAGGCCCTCTTAAGGACAAACATCTCTTCAGTGTTCGTGGTGGAGAA 483
QY 141 MetArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThr 160
Db 484 ATGCTAAAGCTTTTGAGGAAGAATCCACTTTGAACCACTTCACAGGCTGCTACTCACT 543
QY 161 AlaAlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSer 180
Db 544 TCCACAGGAGCTGGATTCATTGACCTTAATCAAGTCTGGGTACAGATCCCTCACTGCT 603
QY 181 GlnTyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr 200
Db 604 CAGTCTCTCGACTATATTTCAGGTTCATGATGATGATGATCTCCATGATCTCAAGAAATGGCTAC 663
QY 201 ThrGlyGluAsnSerProLeuTyrIleTyrProThrAspThrGlySerAsnAlaTyrLeu 220
Db 664 ACTGAGAGAAATAGTCCCTCTATAAATCTCATATGATGATTTGGAAAGAGTCTGATCTC 723
QY 221 AsnValAspTyrValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeuIle 240
Db 724 AATGTGGATTCAATTATTACCTACTTGAAGGACCATGAGGAGCTTCTGAGAAGCTCAT 783
QY 241 ValGlyPheProThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIle 260
```

Db 784 GTGGGATTTCCAGCATATGGTCTATACCTTTATCTCTGAGTGACCTTCTAAGAATGGAATA 843
Qy 261 GlyAlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGlySerGlyIleTrp 280
Db 844 GTGACCTTACTGTTAGTCTGACACACAGGAAAGTACAAATGACAAAGACTCTCTG 903
Qy 281 AlaTyrTyrGluIleCysThrPheLeuLysAsnGlyAlaThrGlnGlyTrpAspAlaPro 300
Db 904 GCTTACTTTGAGATTGTACATTTCTGAATGAAGGAGCCACTGAGATCTTTGATGCCACC 963
Qy 301 GlnGluValProTyrAlaTyrGlnGlyAsnValTrpValGlyTyrAspAsnIleLysSer 320
Db 964 CAGGAAGTACCTATGCTCTATCTGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1023
Qy 321 PheAspIleLysAlaGlnTrpLeuLysHisAsnLysPheGlyGlyAlaMetValTrpAla 340
Db 1024 TTCAGTTGAAGCTCAGTGGCTCAGGACAAATATTAGGAGGCTGCTGCTGCTGCTG 1083
Qy 341 IleAspLeuAspAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSer 360
Db 1084 CTGGACATGGATGACTTCAGTGGTCTTCTCTCTCACCAGGAGCTTCTCTCTGACAACT 1143
Qy 361 ThrLeuLysLysAlaIleGlyLeuGlnSerAlaSerCysThrAla 375
Db 1144 ACTTTAAGAGAGATCTGAATGTACACAGTCAAGTTCGAAGGCC 1188
RESULT 14
LOCUS AY065557 1209 bp mRNA linear ROD 22-MAY-2002
DEFINITION Mus musculus putative secretory protein precursor (Ym2) mRNA,
complete cds.
ACCESSION AY065557
VERSION AY065557.1 GI:18086513
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 1209)
Chang,N.-C.A.
Tissue-specific Expression of Ym2 Protein
JOURNAL Unpublished
2 (bases 1 to 1209)
Chang,N.-C.A.
Direct Submission
TILE
Submitted (05-DEC-2001) Institute of Microbiology & Immunology,
National Yang-Ming University, #155 Sec. 2 Li-Nong St., Beitou,
Taipei 112, Taiwan
FEATURES
Location/Qualifiers
1..1209
/organism="Mus musculus"
/strain="BALB/c"
/db_xref="taxon:10090"
/clone="pGEMT-ez-Ym2.s"
/tissue_type="stomach"
1..1209
/gene="Ym2"
1..1209
/gene="Ym2"
/codon_start=1
/product="putative secretory protein precursor"
/protein_id="AAL57751.1"
/db_xref="GI:18086514"
/translation="MAKILIVTGLAILNLVQLSSGYQLMVCYVTSWAKDRPTEGSPKPG
NIDPCLTHLYAFAMKNEITYLSEQLRDYALNGLDKDRNTLKLTLAIGWKFG
PAPFSWVSTPQNTQTFIKSVIRLQYFNDFGLNDWQYPPGSPGPKDKILFVLVQ
EMRKAPEESTLNHPRILLSTGAGFDIVIKSGYKIPELSGLDYIOVMTVDLHPK
NGYTGENSEPLYSPPDYDKSADLNVDLSIITWKDHGAASEKLIVGFPYAGHTFLLSDP
SKNGIGDPTVSGPPGKTNQGLLAFYEICTFLNEGATEIFDQTEVPYALGNWV
GTDNVSPFKLKAQWLNKNGVWVPLDMDDFSGSFCHQGRFPPLTTTLTKRDLNVHSA
SCKASYRGEL"
326 a 274 c 282 g 327 t
BASE COUNT

ORIGIN

Alignment Scores:
Pred. No.: 1,84e-97 Length: 1209
Score: 1379.00 Matches: 249
Percent Similarity: 80.27% Conservative: 52
Best Local Similarity: 66.40% Mismatches: 74
Query Match: 55.03% Indels: 0
DB: 10 Gaps: 0
US-10-004-219B-14 (1-455) x AY065557 (1-1209)
Qy 1 TyrGlnLeuThrCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPhe 20
Db 64 TACCAGCTGATGTGCTACTATACAGCTGGCTAAGGACAGGCCAACAGAGGGAGTTTC 123
Qy 21 MetProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArg 40
Db 124 AAACCTGGTAATATTGACCCCTGCTGTACTACCTCGATCTATGCCTTTGCTGGGATG 183
Qy 41 GlnAsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrGlnAlaPheAsn 60
Db 184 AAGATATATGATGACACTTACTTAAGTCAGCAGACACTTGGCTGACTATGAAGCATTAAT 243
Qy 61 GlyLeuLysAsnLysAsnSerGlnLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPhe 80
Db 244 GGTCTGAAAGACAGACACACTGAGCTAAATAAACTCTCTGGCCATTGGAGGATGGAAGTTT 303
Qy 81 GlyThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThr 100
Db 304 GGACCTGCCCCGTTTCAGTTCCATGCTCTCTCTCCTCAGAACCCGTCAGACATTTCAATAG 363
Qy 101 SerValIleLysPheLeuArgGlnTrpGluPheAspGlyLeuAspPheAspTrpGluTyr 120
Db 364 TCAGTTATCAGATTCCTTCGTCATATTAACCTTGTATGCGCCCAACCTCGACTGGGAGTAC 423
Qy 121 ProGlySerArgGlySerProGlnAspLysHisLeuPheThrValLeuValGlnGlu 140
Db 424 CCTGGCCCTCGAGGAGGCCCTCTTAAGGACAAACATCTCTTCAGTGTTCCTGGTGAGAA 483
Qy 141 MetArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThr 160
Db 484 ATGCGTAAAGCTTTTGAGGAAGAATCCACTTTTGAACACCATTCCTCAAGGCTGCTACTCAT 543
Qy 161 AlaAlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSer 180
Db 544 TCCACAGAGCTGGATTCATTCAGTAATCAAGCTGGGTACAGATCCCTCAAGCTGCT 603
Qy 181 GlnTyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr 200
Db 604 CAGTCTCTCGACTATATTCAGGTCTATGACATATGATCTCCATGATCCTTAAGAATGCTTAC 663
Qy 201 ThrGlyGluAsnSerProLeuTyrIleTyrProThrAspThrGlySerAsnAlaTyrLeu 220
Db 664 ACTGGAGAAAATAGTCCCTCTATAAATCTCATATGATGATGAGAGAGTGTCTGATCTC 723
Qy 221 AsnValAspTyrValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeuIle 240
Db 724 AATGGGATTCAAATTTATTCCTCTGGAAGGACCATGAGGAGCTTCTTGAGAGGCTCAT 783
Qy 241 ValGlyPheProThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIle 260
Db 784 GTGGATTTCCAGCATATGGTCTATACCTTTATCTCTGAGTGACCTTCTTAAGAATGGAATA 843
Qy 261 GlyAlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGlySerGlyIleTrp 280
Db 844 GGTGACCTTACTGTTAGTCTGACACACAGGAAAGTACAAATGACAAAGACTCTCTG 903
Qy 281 AlaTyrTyrGluIleCysThrPheLeuLysAsnGlyAlaThrGlnGlyTrpAspAlaPro 300
Db 904 GCTTACTTTGAGATTGTGACATTTCTGGAATGAAGGAGCCACTGAGATCTTTGATGCCACC 963
Qy 301 GlnGluValProTyrAlaTyrGlnGlyAsnValTrpValGlyTyrAspAsnIleLysSer 320

repeat_region /note="L2 repeat: matches 2595. .2748 of consensus"
37599. .37717
/note="MER94 repeat: matches 1. .119 of consensus"
37777. .38064
/note="L2 repeat: matches 2188. .2502 of consensus"
38106. .38203
/note="MIR repeat: matches 42. .140 of consensus"
39060. .39162
/note="MIR repeat: matches 77. .192 of consensus"
39248. .39475
/note="MIR repeat: matches 17. .261 of consensus"
40505. .40582
/note="L2 repeat: matches 2661. .2740 of consensus"
42696. .42998
/note="Alusg repeat: matches 1. .301 of consensus"
43366. .45787
/note="TRIGER1 repeat: matches 1. .2418 of consensus"
46715. .46908
/note="MIR repeat: matches 1. .206 of consensus"
48333. .49296
/note="LTR5 repeat: matches 1. .968 of consensus"
49597. .49909
/note="Alusg repeat: matches 1. .313 of consensus"
49918. .49957
/note="20 copies 2 mer tt 77% conserved"
49962. .50026
/note="L1P47 repeat: matches 6082. .6145 of consensus"
50027. .50995
/note="LTR5 repeat: matches 21. .969 of consensus"
50996. .51523
/note="L1P47 repeat: matches 5543. .6082 of consensus"
52579. .52919
/note="L2 repeat: matches 2344. .2708 of consensus"
53182. .53681
/note="LTR10D repeat: matches 1. .513 of consensus"
56386. .56623
/note="MIR repeat: matches 2. .215 of consensus"
58241. .58272
/note="16 copies 2 mer aa 84% conserved"
58587. .58978
/note="WSTA repeat: matches 2. .426 of consensus"
59106. .59464
/note="MLTIF repeat: matches 163. .539 of consensus"
59514. .59561
/note="L2 repeat: matches 2576. .2623 of consensus"
59683. .59750
/note="L2 repeat: matches 2635. .2734 of consensus"
60622. .62283
/note="L1M1 repeat: matches -1390. .639 of consensus"
62284. .62605
/note="Alusx repeat: matches 1. .312 of consensus"
62606. .63401
/note="L1M1 repeat: matches 639. .1409 of consensus"
63397. .63569
/note="L1M2 repeat: matches 2559. .2735 of consensus"
63655. .66952
/note="L1M2 repeat: matches 2926. .6308 of consensus"
67388. .67655
/note="L1P83 repeat: matches 5860. .6150 of consensus"
67741. .68108
/note="L1M4 repeat: matches 5483. .5856 of consensus"
68311. .68618
/note="AluY repeat: matches 1. .311 of consensus"
73419. .73525
/note="L2 repeat: matches 2591. .2691 of consensus"
73535. .73690
/note="L2 repeat: matches 2570. .2726 of consensus"
73775. .74035
/note="L2 repeat: matches 2425. .2705 of consensus"
74535. .74690
/note="MIR repeat: matches 46. .205 of consensus"

Pred. No.: 7,47e-94 Length: 91569
Score: 1362.50 Matches: 317
Percent Similarity: 34.12% Conservative: 3
Best Local Similarity: 33.80% Mismatches: 6
Query Match: 54.37% Indels: 613
DB: 9 Gaps: 5
US-10-004-219b-14 (1-455) x AL356387 (1-91569)
Qy 130 AsplysHisLeuPheThrValLeuValGlnGluMetArgGluAlaPheGluGlnGluAla 149
Db 6262 GATGTCCTATTATGCTTATTATCTGTAGGAATGCGTAGGCTTTTGAGCAGAGGCC 6321
Qy 150 LysGlnIleAsnLysProArgLeuMetValThrAlaAlaValAlaAlaGlyIleSerAsn 169
Db 6322 AAGCAGATCAACAGCCAGCGCTGATGTCATCTGTCAGTAGCTGCTGGCATCTCCAAT 6381
Qy 170 IleGlnSerGlyTyrGluIleProGlnLeuSerGln 181
Db 6382 ATCCAGTCTGGCTATGAGATCCCCCACTGTACAGTGAGTGATGTGCTTATCTTCAA 6441
Qy 181 181
Db 6442 CTCCTGGAGGTGTCACTCGGGCACACTAGACTTGTCTTGGTGTGGCTTATTCAGGG 6501
Qy 181 181
Db 6502 ACCTTGTTTAGGAGACTTTAGNAGTGGTGCTCTGTGCTGCATAGGTCTGGGAATGACC 6561
Qy 182 182
Db 6562 ATCAGATGATTTTAAATCTCCACCCACCTCAGGTACCTGGAGTACATCATCTCATG 6621
Qy 190 ThrTyrAspLeuHisGlySerTyrGluGlyTyrThrGlyGluAsnSerProLeuTyrLys 209
Db 6622 ACCTACGACCTCCATGGCTCTCTGGAGGGCTACACTGAGAGAACAGCCCCCTTACAAA 6681
Qy 210 TyrProThrAspThrGlySerAsnAlaTyrLeuAsnVal 222
Db 6682 TACCGAGTACACCGGCGACACGCTTACCTCAATGT -GGTGAGTCCCTGTGACAGATGC 6740
Qy 222 222
Db 6801 AAGTTTAGTGGTTTCTCAAATGGGAATTAGGTGGCCATTAATTAATAATATTCTCAT 6860
Qy 222 222
Db 6861 TTAATTTACACAGAGACTGGGACACAAAGTAACATAAATTTCTCAAGTGTCTCACTTA 6920
Qy 222 222
Db 6921 AACAACTGTCTCAAGTTGATCGTATTTCATTAACACTACAACCTTCCACAAAATGTTTGA 6980
Qy 222 222
Db 6981 AAAAATACTAATCCATCTGGAAATTAACAGCATAGAGTTTTTACCTGTCGGAATAGGGACT 7040
Qy 222 222
Db 7041 AATATACTAAGTACTGGGTCTCAGCTGTGGGCCATGTAACTAACCCACTGACATTG 7100
Qy 223 223
Db 7101 CAGGATTATGTATCATGAACACTACTGGAAGCAATGAGCACCAGCTGAGAAGCTCATCGTT 7160
Qy 242 GlyPheProThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIleGly 261
Db 7161 GGATTCCTTACCTATGGACACAACACTTTCATCTGTGAGCAACCCCTCCACACTGGAATGGT 7220

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame plus p2n model

Run on: June 29, 2003, 21:02:46 ; Search time 249.712 Seconds
(without alignments)
4103.372 Million cell updates/sec

Title: US-10-004-219B-14

Perfect score:

Sequence: 1 YQLTCYFTNWAQYRPGGLRF.....QQNCOAGLVFDTSCDCCNWA 455

Scoring table:	BLOSUM62		
	Xgapop	10.0	Xgapext 0.5
	Ygapop	10.0	Ygapext 0.5
	Fgapop	6.0	Fgapext 7.0
	Delop	6.0	Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Minimum DB seq	length: 0
Maximum DB seq	length: 2000000000

Post-processing: Minimum March 0%

Post-processing: Minimum Match 0%
Maximum Match 100%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

```

Command line parameters:
-MODEL=framed-pzn.model -DEV=xlp
-Q=Cg2n1/USPTO.spool/US10004219/runat.24062003.160227.8151/app_query.fasta_1.1877
-D=N Genseq 101002 -GFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPL=0
-DB=N Genseq 101002 -GFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPL=0
-LOOPT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10004219 @CGN 1 1.626 @runat.24062003.160227.8151 -NCPU=6 -ICPU=3
-NO MWAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCGAPOP=6
-FCGAPT=7 -VCAPOP=10 -VGAEXT=0.5 -DELOP=6 -FCGAPT=7

```

Database : N_Geneseq_101002:*

1:	/SID52/cgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2:	/SID52/cgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3:	/SID52/cgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4:	/SID52/cgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5:	/SID52/cgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6:	/SID52/cgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7:	/SID52/cgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8:	/SID52/cgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9:	/SID52/cgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10:	/SID52/cgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11:	/SID52/cgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12:	/SID52/cgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13:	/SID52/cgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14:	/SID52/cgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15:	/SID52/cgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16:	/SID52/cgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17:	/SID52/cgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18:	/SID52/cgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19:	/SID52/cgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20:	/SID52/cgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21:	/SID52/cgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22:	/SID52/cgdata/geneseq/geneseqn-emb1/NA2001A.DAT
23:	/SID52/cgdata/geneseq/geneseqn-emb1/NA2001B.DAT
24:	/SID52/cgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query			ID	Description
		Match	Length	DB		
1	2506	100.0	1368	22	AAH42013	Disease treatment
2	2506	100.0	1678	22	AAH42025	Disease treatment
3	2137	85.3	1369	24	ABN59993	Novel human coding
4	1415	56.5	1469	22	AAH42023	Disease treatment
5	1323	52.8	1636	19	AAH10435	Human chitinase c1
6	1323	52.8	1636	22	AAH21847	MO-218 clone of hu
7	1323	52.8	1636	22	AAH03759	Human chitinase cD
8	1323	52.8	1636	24	ABL57380	Human chitinase cD
9	1323	52.8	1643	18	AAH50833	Human chitinase cD
10	1317	52.6	1656	19	AAH10436	Human chitinase c1
11	1317	52.6	1656	20	AAH21848	Human chitinase c1
12	1317	52.6	1656	22	AAH03760	Human chitinase cD
13	1317	52.6	1656	24	ABL57381	Human chitotriosid
14	1316	52.5	1768	18	AAH89181	Human chitotriosid
15	1311.5	52.3	1713	18	AAH50834	Human chitinase cD
16	1297	51.8	1637	18	AAH89180	Human chitotriosid
17	1011.5	40.4	2504	16	AAQ90443	Murine oviduct spe
18	1008.5	40.2	2346	23	ABL03597	Drosophila melanog
19	997	39.8	638	24	ABK11713	DNA encoding novel
20	995	39.7	746	22	AAH88635	Human digestive sy
21	978.5	39.0	1594	19	AAH21688	DNA encoding a hum
22	970.5	38.7	1432	19	AAH21689	DNA encoding a hum
23	970.5	38.7	1433	19	AAH13925	Human cartilage gp
24	970.5	38.7	1496	18	AAH97127	Human cartilage gp
25	970.5	38.7	1526	19	AAH13926	Human cartilage gp
26	970.5	38.7	1594	19	AAH21687	DNA encoding a hum
27	963.5	38.4	1152	18	AAH99445	Bovine whey protei
28	962.5	38.4	1681	16	AAH08542	YKL-40 gene. Homo
29	962.5	38.4	1681	21	AAZ94901	Human cancer marke
30	952.5	38.0	2366	16	AAQ90444	Hamster oviduct sp
31	943.5	37.6	1994	16	AAQ90442	Bovine oviduct spe
32	936	37.4	17368	23	ABL03596	Drosophila melanog
33	923.5	36.9	1391	24	AAH62615	cDNA sequence #402
34	890.5	35.5	13497	23	ABL03557	Drosophila melanog
35	873.5	34.9	1474	22	AAH23078	Osteoarthritis tis
36	863	34.4	2452	18	AAH22557	Manduca sexta larv
37	853	34.0	1433	23	ABL15099	Drosophila melanog
38	832.5	33.2	3552	23	ABL15098	Drosophila melanog
39	798	31.8	16489	23	ABL03556	Drosophila melanog
40	789.5	31.5	2510	23	ABL06628	Drosophila melanog
41	768.5	30.7	2089	23	ABL29841	Drosophila melanog
42	760	30.3	966	9	AAH1756	Gene encoding poly
43	723.5	28.9	1302	23	ABL19745	Drosophila melanog
44	723.5	28.9	3302	23	ABL19744	Drosophila melanog
45	715.5	28.6	3751	23	ABL19164	Drosophila melanog

ALIGNMENTS

RESULT 1	
AAH42013	
ID	AAH42013 standard; DNA; 1368 BP.
XX	
XX	
XX	AAH42013;
XX	
XX	
XX	24-AUG-2001 (first entry)
XX	
XX	
XX	Disease treatment related oligonucleotide SEQ ID NO: 3.
XX	
XX	
KW	Disease treatment; infection; chronic occlusive pulmonary disease;
KW	bronchial asthma; ds.
XX	
OS	Homo sapiens.
XX	
XX	
FN	WO200136633-Al.
XX	
XX	
PD	25-MAY-2001.

XX
PF
14-NOV-2000; 2000WO-JP08015.

XX
PR 15-NOV-1999; 99JP-0324467.

XX
PA (TAKE) TAKEDA CHEM IND LTD.

PI Nakanishi A. Morita S:

11 NURENBERG II, MOTTERS
 XX
 DR WPI: 2001-397791/42.

DA
XX
PT
PT

New proteins, peptides and DNA for treatment of bronchial asthma,
chronic occlusive lung disease and infectious disease -

XX
PS Claim 5; Page 100; 114pp; Japanese.

XX CC The present invention provides the sequence of a protein which can be
CC used in the treatment and prevention of infectious diseases. Inhibitors
CC of the protein can be used to treat bronchial asthma and chronic
CC occlusive pulmonary disease. The present sequence is an oligonucleotide
CC described in the exemplification of the invention.

Sequence 1368 BP; 324 A; 398 C; 349 G; 297 T; 0 other;
XX

Alignment Scores:

Pred. No.:	8,966-220	Length:	1368
Score:	2506.00%	Matches:	455
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	22	Gaps:	0

US-10-004-219B-14 (1-455) x AAH42013 (1-1368)

Qy	1	TyrGlnLeuThrCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPhe	20
Db	1	TACCAGCTGACATGCTACTTCAACCACTGGGCCCAAGTACCGGCCAGGCGCTTC	60
Qy	21	MetProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArg	40
Db	61	ATGCCCTGCAACATCGACCCCTGCTGTACCACCTGATCTACGGCTTTGTGGGAGG	120
Qy	41	GlnAsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrGlnAlaPheAsn	60
Db	121	CAGAACACAGAGATCACCACCATCAATGGATGATGTGACTCTCTACCAAGCTTTCAT	180
Qy	61	GlyLeuLysAsnLysAsnSerGlnLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPhe	80
Db	181	GGCCTGAAAAATAAGAACAGCCAGCTGAAAACTCTCTGGCCATTGGAGGCTGAACTTC	240
Qy	81	GlyThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThr	100
Db	241	GGAGCTGCCCTTTTCACTGCCATGTTTCTACTCTCTGAGAACCCGACAGCTTTCATCAC	300
Qy	101	SerValIleLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyr	120
Db	301	TCAGTCATCAAAATTCCTGGCCAGTATGAGTTTGACGGGCTGGACTTTTGACTGGGAGTAC	360
Qy	121	ProGlySerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValGlnGlu	140
Db	361	CCTGGCTCTCGTGGGAGCCCTCTCAGGACAAGCATCTCTTCACTGTCTGGTGCAGGAA	420
Qy	141	MetArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThr	160
Db	421	ATGCGGTGAAGCTTTTGGACGAGGAGCCCAAGCAGATCAACAAGCCAGGCTGATGGTCACT	480
Qy	161	AlaAlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSer	180
Db	481	GCTGCAGTAGCTGCTGGCATCTCCAAATATCAGTCTGGCTATGAGATCCCCCAACTGTCA	540
Qy	181	GlnTyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr	200
Db	541	CAGTACCTGGATACCATCATGTCTATGACTTACCACTTCCATGGCTCCTGGGAGGGCTAC	600

Qy	201	ThrGlyGluAsnSerProLeuTyfYrLyfYrProThrAspThrGlySerAsnAlaTyfLeu	220
Db	601	ACTGGAGAGAACAGCCCCCTCTCAAAATACCGGACTGACACCGGACGACACGCCTACCTC	660
Qy	221	AsnValAspTyfValMetAsnTyfTrpLysAspAsnGlyAlaProAlaGluLysLeuIle	240
Db	661	AATGTGATTTATGTCATGAACACTCTGGAAGACAAATGGAGCACAGCTGAGAAGCTCATC	720
Qy	241	ValGlyPheProThrTyfGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIle	260
Db	721	GTTCGATTCCCTACCTATGGACCAACTTCATCCTGAGCAACCCCTCCACACCTGGAAATT	780
Qy	261	GlyAlaProThrSerGlyAlaGlyProAlaGlyProTyfValLysGluSerGlyIleTrp	280
Db	781	GGTGCCTCCCACTCTGTGTGTCCTGCTGGGCCCTATGCCAAGGAGTCTGGGATCTGG	840
Qy	281	AlaTyfTyfGluIleCysThrPheLeuLysAsnGlyAlaThrGlnGlyTrpAspAlaPro	300
Db	841	GCATTACTACGAGATCTGTACTCTTCGTAAATAATGGAGCCACTCAGGATGGGATGCCCT	900
Qy	301	GlnGluValProTyfValTyfGlnGlyAsnValTrpValGlyTyfAspAsnIleLysSer	320
Db	901	CAGGAAGTGCTTATGGCTATCAGGGCAATGTGTGGGTGGCTATGACACATCAAGAGC	960
Qy	321	PheAspIleLysAlaGlnTrpLeuLysHisAsnLysPheGlyGlyAlaMetValTrpAla	340
Db	961	TTCCGATATTAAAGCTCAATGGCTTAAGCAACAATAATTTGGAGCGGCATGGTCTGGGCC	1020
Qy	341	IleAspLeuAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSer	360
Db	1021	ATTGATCTGGATGACTTCACTGGACCTTCTGCAACGAGGGCAAGTTTCCCTTAATCTCC	1080
Qy	361	ThrLeuLysLysAlaLeuGlyLeuGlnSerAlaSerCysThrAlaProAlaGlnProIle	380
Db	1081	ACCTGAGAGAGGCCCTTCGCCCTCGAGTGCAGTTGACGGCTCCAGCTCAGCCCATTT	1140
Qy	381	GluProIleThrAlaAlaProSerGlySerGlyAsnGlySerGlySerSerSerGly	400
Db	1141	GAGCCAATAACTGCTGCTCCAGTGGCAGCGGAAACGGGAGCGGAGTAGCAGCTCTGGA	1200
Qy	401	GlySerSerGlyGlySerGlyPheCysAlaValArgAlaAsnGlyLeuTyfProValAla	420
Db	1201	GGCAGCTCGGGAGGGCAGTGGATTCTGTGCTGTGAGAGCCAACGGCTCTTACCCCGTGGCA	1260
Qy	421	AsnAsnArgAsnAlaPheTrpHisCysValAsnGlyValThrTyfGlnGlnAsnCysGln	440
Db	1261	AATPACAGAAATGCCCTTCGGCACTGGGTGAATGGAGTCACTACACAGCAGAACTGCCAG	1320
Qy	441	AlaGlyLeuValPheAspThrSerCysAspCysAsnTrpAla	455
Db	1321	GCCGGGCTTGCTTTCACACACAGCTGATTGCTGCTGCAACTGGGCA	1365

RESULT 2

AAH42025

ID AAH42025 standard; DNA; 1678 BP.

XX

AC AAH42025;

XX	
XX	
DT	24-AUG-2001 (first entry)
XX	
XX	
DE	Disease treatment related oligonucleotide SEQ ID NO: 16.
XX	
KW	Disease treatment; infection; chronic occlusive pulmonary disease;
KW	bronchial asthma; ds.

OS Homo sapiens.

XXIX

PN WO200136633-

XX
25 MAY 2001

PD YY

XX 15-NOV-1999; 99JP-0324467.

XX (TAKE) TAKEDA CHEM IND LTD.

XX Nakanishi A, Morita S;

XX WPI; 2001-397791/42.

XX New proteins, peptides and DNA for treatment of bronchial asthma,
PT chronic occlusive lung disease and infectious disease

XX Example 5; Page 107-108; 114pp; Japanese.

XX The present invention provides the sequence of a protein which can be
CC used in the treatment and prevention of infectious diseases. Inhibitors
CC of the protein can be used to treat bronchial asthma and chronic
CC occlusive pulmonary disease. The present sequence is an oligonucleotide
CC described in the exemplification of the invention.

XX Sequence 1678 BP; 397 A; 497 C; 418 G; 366 T; 0 other;

Alignment Scores:

Pred. No.: 1.19e-219 Length: 1678
Score: 2506.00 Matches: 455
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 22 Gaps: 0

US-10-004-219b-14 (1-455) x AAH42025 (1-1678)

QY 1 TyrGlnLeuThrCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPhe 20
DB 206 TACCAGCTGACATGCTTACCTACCACTGGGCCAGTACCGCCAGGCTGGGGCCCTTC 265
QY 21 MetProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArg 40
DB 266 ATGCCTGCAACATCATGACCCCTGCTCTGACCACCTGATCTAGCCCTTTGCTGGGAGG 325
QY 41 GlnAsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrGlnAlaPheAsn 60
DB 326 CAGAACACAGAGATCAACACCATCGAATGGAATGATGTGACTCTCTACCAAGCTTTCAAT 385
QY 61 GlyLeuLysAsnLysAsnSerGlnLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPhe 80
DB 386 GGCTGAAATAAAGAACACCCAGCTGAAACTCTCTGGCCATTGGAGCTGGAACTTC 445
QY 81 GlyThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThr 100
DB 446 GGGAGTGGCCCTTTTCACTGCCATGGTTTCTACTCTGAGAACCGCCAGACTTTTCATCAC 505
QY 101 SerValIleLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyr 120
DB 506 TCAGTCATCAAAATTCCTGGCCAGTATGATTTGACGGGCTGGACTTTGACTGGGAGTAC 565
QY 121 ProGlySerArgGlySerProGlnAspLysPheThrValLeuValGlnGlu 140
DB 566 CTGGCTCTCGTGGAGCCCTCTCTAGGACAGCATCTCTACTGTCTCTGGTGGAGAA 625
QY 141 MetArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThr 160
DB 626 ATGGGTGAAGCTTTTTCAGCAGGAGGCGCAAGCAGATCAACAGCCAGGCTGATGGTCACT 685
QY 161 AlaAlaValAlaAlaGlyLysSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSer 180
DB 686 GCTGCAGTAGTCTGGCATCTCCATATATCCAGTCTGGCTATGAGATCCCCCAACTGTCA 745
QY 181 GlnTyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr 200
DB 746 CAGTACTCTGGACTACATCCATGTCATGACCTACGACCTCCATGGCTCCCTGGGAGGCTAC 805
QY 201 ThrGlyGluAsnSerProLeuTyrLysTyrProThrAspThrGlySerAsnAlaTyrLeu 220

DB 806 ACTGGAGAGAACAGCCCTCTTACAAATACCGACTGACCCGGCAGCAACCCCTACCTC 865
QY 221 AsnValAspTyrValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeuIle 240
DB 866 AATGTGGATTATGTGTAATCTACTGGAAGGAGCAATGGAGCACCAGCTGAGAAAGCTCATC 925
QY 241 ValGlyPheProThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIle 260
DB 926 GTTGGATTCCCTTACCTATGGACACAACTTCTCTGAGCAACCCCTCCAAACACTGGAATT 985
QY 261 GlyAlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyIleTrp 280
DB 986 GGTGCCCCCACCCTCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1045
QY 281 AlaTyrTyrGluIleCysThrPheLeuLysAsnGlyAlaThrGlnGlyTrpAspAlaPro 300
DB 1046 GCTTACTACGAGATCTGTACCTTCTGAAATAATGAGCCACCTCAGGGATGGATGCCCT 1105
QY 301 GlnGluValProTyrAlaTyrGlnGlyAsnValTrpValGlyTyrAspAsnIleLysSer 320
DB 1106 CAGGAAGTGCCTTATGCTATCAGGGCAATGTGTGGTGGCTATGACAAACATCAAGAGC 1165
QY 321 PheAspIleLysAlaGlnTrpLeuLysHisAsnLysPheGlyGlyAlaMetValTrpAla 340
DB 1166 TTCCGATATTAAAGGCTCAATGGCTTAAGCACACAAATTTGGAGGCGCCATGCTCTGGGCC 1225
QY 341 IleAspLeuAspAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSer 360
DB 1226 ATTGATCTGGATGACTTCTACCTGGCACTTTCTGCAACCCAGGGCAAGTTTCCCTAATCTCC 1285
QY 361 ThrLeuLysLysAlaLeuGlyLeuGlnSerAlaSerCysThrAlaProAlaGlnProIle 380
DB 1286 ACCCTGAGAGAGCCCTCGGCTGCGAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAG 1345
QY 381 GluProIleThrAlaAlaProSerGlySerGlyAsnGlySerGlySerSerSerSerGly 400
DB 1346 GAGCCAAATACTGCTGCCAGTGGCAGCGGAAACCGGAGCGGAGTAGCAGCTCTGGA 1405
QY 401 GlySerSerGlyGlySerGlyPheCysAlaValArgAlaAsnGlyLeuTyrProValAla 420
DB 1406 GGCAGCTGGGAGGAGTGGATCTGTGCTGTCAGAGCAACGGCTCTTACCCCTGGGCA 1465
QY 421 AsnAsnArgAsnAlaPheTrpHisCysValAsnGlyValThrTyrGlnGlnAsnCysGln 440
DB 1466 AATAACAGAAATGCCCTTCTGGCACTGCGTGAATGGAGTCACTACACAGCAAGTCCAG 1525
QY 441 AlaGlyLeuValPheAspThrSerCysAspCysCysAsnTrpAla 455
DB 1526 GCCGGGCTTGTCTTCGACACACAGCTGTGATTGCTGCAACTGGGCA 1570

RESULT 3

ABN59993
ID ABN59993 standard; cDNA; 1369 BP.
XX AC ABN59993;
XX XX 28-JUN-2002 (first entry)
XX DE Novel human coding sequence SEQ ID NO: 404.
XX KW Human; antianemic; vulnerary; antiinflammatory; immunomodulator;
XX KW antinfertility; cerebroprotective; cytostatic; rheumatic; gene therapy;
XX KW neuroprotective; antiparkinsonian; protein therapy; EST;
XX XX expressed sequence tag; gene; ss.
XX OS Homo sapiens.
XX XX WO200222660-A2.
XX XX 21-MAR-2002.
XX PD 10-SEP-2001; 2001WO-US26015.
XX PF

Claim 3; Page 38-40; 63pp; English.

This sequence encodes a novel human chitinase isolated from clone MO-218. Chitinases are useful for treating or preventing fungal infection and as immunogens for generating antibodies which are used to purify, detect and quantify chitinases, e.g. for diagnosis of Gaucher's disease. The nucleic acid sequence of the chitinase is also useful as a probe to identify and isolate genomic DNA encoding chitinases or similar proteins, or cells expressing them or to generate transgenic ('knockout') rodents. It can also be used in hybridisation assays and to detect genetic alterations in the chitinase gene related to disease. Agents that inhibit this protein may be useful in treatment of Gaucher's disease and rheumatoid arthritis, where overexpression of the protein can damage the extracellular matrix. Chitinase also improves the activity of other antifungal agents and may allow a reduction in the dose of such agents, and thus of their side effects.

Sequence 1636 BP; 361 A; 491 C; 440 G; 344 T; 0 other;

Alignment Scores:

Pred. No.:	2,356-111	Length:	1636
Score:	1323.00	Matches:	240
Percent Similarity:	67.39%	Conservative:	70
Best Local Similarity:	52.17%	Mismatches:	126
Query Match:	52.79%	Indels:	24
DB:	19	Gaps:	3

US-10-004-219b-14 (1-455) x AAV10435 (1-1636)

```

QY 2 GlnLeuThrCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPheMet 21
DB 68 AACTGGTCTGCTACTTCAACAACTGGGCCAGTACAGACAGGAGGAGGCTGCTTCCTG 127

QY 22 ProAspAsnIleAspProCysLeuCysThrHisLeuIleThrAlaPheAlaGlyArgGln 41
DB 128 CCAAGAGACTTGGACCCAGCCCTTTGCACCACCTCATCTACGCCCTTGGCTGGCATGACC 187

QY 42 AsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrGlnAlaPheAsnGly 61
DB 188 AACCCAGCTGAGCACCCTGAGTGGATGACGAGACTCTTACAGGAGTTCATATGCG 247

QY 62 LeuIleAsnLysAsnSerGlnLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPheGly 81
DB 248 CTGAAGAAGATGAATCCCAAGCTGAAGACCCCTTTAGCCATCGAGGCTGGAATTTGGC 307

QY 82 ThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThrSer 101
DB 308 ACTCAGAAGTTCAGATATGGTAGCCAGCCGCCCAACCCGTCAGACCTTTGTCAACTCG 367

QY 102 ValIleLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyrPro 121
DB 368 GCCATCAGGTTCTCGCAAAATACAGCTTTGACGGCTTGACCTTGCTGGGATACCCA 427

QY 122 GlySerArgGlySerProGlnAspLysHisLeuPheThrValLeuValGlnGluMet 141
DB 428 GGAAGCCAGGGAGCCCTGCGTAGAACAGGAGCGCTTCACACCCCTGGTGTACAGACTTG 487

QY 142 ArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThrAla 161
DB 488 GCCAATGCCCTTCAGCAGGAGGCCAGCCCTCAGGGAAGAACGCTTCTTCTGTAGTGCA 547

QY 162 AlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSerGln 181
DB 548 GCGGTTCCAGCTGGCAGACCTATGTGGATGCTGGATACGAGGTGGACAAATCGCCAG 607

QY 182 TyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThr 201
DB 608 AACCTGGATTTTGTCAACCTTATGGCTACGACTTCATGGCTCTTGGGAGAGGTCAGC 667

QY 202 GlyGluAsnSerProLeuTyrLysTyrProThrAspThrGlySerAsnAlaTyrLeuAsn 221
DB 668 GGACATAACAGCCCTCTCAAGAGGCGCAAGAGAGAGTGGTGCAGCAGCCAGCTCAAC 727

```

```

QY 222 ValAspTyrValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeuIleVal 241
DB 728 GTGGATGCTGCTGTCAACAGAGTGGCTGCAGAGGGGACCCCTGCCAGCAAGCTGATCCTT 787

QY 242 GlyPheProThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIleGly 261
DB 788 GGCATGCCCTTACTACGGACGCTCTCTACACCTGGCTCTCTATCAGACACCAAGAGTGGG 847

QY 262 AlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyIleTrpAla 281
DB 848 GCCCAGCAGCAGGCTTGGCACTCCAGCCCTTCACCAGGAAGAGGAGGATGTGGCC 907

QY 282 TyrTyrGluIleCysThrPheLeuLysAsnGlyAlaThrGlnGlyTrpAspAlaProGln 301
DB 908 TACTATGAAGTCTGCTCTGG-----AAGGGGGCCACCACAGAGAATCAGGATCAG 961

QY 302 GluValProTyrAlaTyrGlnGlyAsnValTrpValGlyTyrAspAsnIleLysSerPhe 321
DB 962 AAGGTGCCCTTACATCTTCGGGACCAACAGTGGTGGGCTTTGTATGATGTGGAGAGCTTC 1021

QY 322 AspIleLysAlaGlnTrpLeuLysHisAsnLysPheGlyGlyAlaMetValTrpAlaIle 341
DB 1022 AAAACCAAGCTCAGCTATCTGAGCAGAGGAGCTGGGGGGGCCATGCTCTGGGCACGTG 1081

QY 342 AspLeuAspAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSerThr 361
DB 1082 GACTTAGATGACTTTGCCGGCTTCTCTGCAACAGGCGCGATACCCCTCATCCAGACG 1141

QY 362 LeuLysLysAlaLeu-----GlyLeuGlnSerAlaSerCysThr 374
DB 1142 CTACGGCAGGAAGTCTGCTTCCATCTTCCAGCAGCCACCCAGAGCTTGAAGTTCCA 1201

QY 375 AlaProAlaGlnProIleGluProIleThrAlaAlaProSerGlySerGlyAsnGlySer 394
DB 1202 AACCAAGCTCAGCCCTCTGAACCT----- 1225

QY 395 GlySerSerSerSerGlyGlySerSerGlyGlySerGlyPheCysAlaValArgAlaAsn 414
DB 1226 -----GAGCATGGCCCGCCAGCCCTGGACAAGACACGTTCTGCCAGGGCAAGCTGAT 1276

QY 415 GlyLeuTyrProValAlaAsnAsnArgAsnAlaPheTrpHisCysValAsnGlyValThr 434
DB 1277 GGGCTCTATCCCAATCTCTGGGAACGGTCCAGCTTCTACAGCTGTGCAGCGGGCGGCTG 1336

QY 435 TyrGlnGlnAsnCysGlnAlaGlyLeuValPheAspThrSerCysAspCysCysAsnTrp 454
DB 1337 TTCCAGCAAGCTGCCAGCAGGCTGTGTTCAGCAACTCTCTGCAATGCTGCACCTGG 1396

RESULT 6
AAZ21847
ID AAZ21847 standard; DNA; 1636 BP.
XX AAZ21847;
AC AAZ21847;
XX 10-DEC-1999 (first entry)
XX DT
XX MO-218 clone of human Chitinase, with noncoding 5'/3' regions.
XX DE
XX chitin; fungal infection; immunocompromised; AIDS; chemotherapy;
XX organ transplant; parasite; chitin-binding; allele; vector;
XX truncated protein; ds.
XX Homo sapiens.
XX Key
XX Location/Qualifiers
XX CDS
XX 2..1402
XX /*tag= a
XX /product= Human_Chitinase
XX sig_peptide
XX 2..65
XX /*tag= b
XX /product= Signal_peptide
XX 66..1402
XX mat_peptide
XX /*tag= c

```

/product= Mature_protein

FT XX W09946390-A1.
 PN XX 16-SEP-1999.
 PD XX 12-MAR-1999; 99WO-US05343.
 PF XX 12-MAR-1998; 98US-0039198.
 PR XX (ICOS-) ICOS CORP.
 PA XX
 PI XX Gray PW, Tjoelker LW;
 XX WPI; 1999-551417/46.
 DR XX P-PSDB; AAY42425.
 DR XX
 XX Novel chitin-binding fragments of human chitinase used to treat fungal
 PT infections in animals
 PT
 PS Example 1; Page 55-57; 83pp; English.
 CC This is the nucleotide sequence of an allelic form of the human
 CC chitinase enzyme, which is capable of degrading Chitin (a linear
 CC homo polymer of beta-1,4-linked N-acetylglucosamine residues).
 CC Chitinase fragments can be used to screen for proteins or other
 CC molecules that specifically bind to the chitin-binding domain of human
 CC chitinase or that modulate its activity. These compounds are useful for
 CC immunization, as well as for purifying chitinase, as well as for
 CC detection and quantification of chitinase. Polynucleotide fragments of
 CC the invention are useful as a source of probes and primers, and to
 CC express the proteins recombinantly. The chitinase fragments, when
 CC conjugated to antifungal compounds, are used to treat animals,
 CC especially humans, infected with chitin-containing parasites such as
 CC fungi. Fungal infection treated include candidiasis, aspergillosis,
 CC coccidioidomycosis, blastomycosis, paracoccidioidomycosis,
 CC mucormycosis, histoplasmosis, cryptococcosis, chromoblastomycosis,
 CC sporotrichosis, and dermatophytoses.
 CC Chitin can be degraded by the enzyme chitinase. Use of whole chitinase
 CC protein for treating infections, especially fungal infections, is
 CC problematic. In view of the increasing incidents of life-threatening
 CC fungal infection in e.g. immunocompromised individuals, there exists a
 CC need for identifying new compounds for treating fungal infection. The
 CC chitin-binding fragments of the present invention provide this need.
 XX
 SQ Sequence 1636 BP; 361 A; 491 C; 440 G; 344 T; 0 other;

Alignment Scores:

Pred. No.: 2,35e-111 Length: 1636
 Score: 1323.00 Matches: 240
 Percent Similarity: 67.39% Conservative: 70
 Best Local Similarity: 52.17% Mismatches: 126
 Query Match: 52.79% Indels: 24
 DB: 20 Gaps: 3

US-10-004-219B-14 (1-455) x AA221847 (1-1636)

QY 2 GlnLeuThrCysThrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPheMet 21
 DB 68 AAACGTGGTCTGCTACTTACCAACTGGGCCAGTAGACAGACAGGGGAGGCTCGCTTCCTG 127
 QY 22 ProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArgGln 41
 DB 128 CCNAGGACTTGGACCCAGCCCTTTGCACCCACCTCATCTAGCCCTTCGTGGCATGCC 187
 QY 42 AsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrGlnAlaPheAsnGly 61
 DB 188 AACCCACGACTGAGCACCAGCTGAGTGGATGACGAGACTCTTACCCAGGATTCATGGC 247
 QY 62 LeuLeuAsnLeuAsnSerGlnLeuIleThrLeuLeuAlaIleGlyTyrAsnPheGly 81
 DB 248 CTGAAGAAGATGAATCCCAAGCTGAAGACCCCTGTTAGGCCATCCGAGGCTGGAATTCGCGC 307

QY 82 ThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThrSer 101
 DB 308 ACTCAGAAGTTTCACAGATATGCTAGCCACGCCCAACACCTGACACCTTTGTCACACTCG 367
 QY 102 ValIleLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyrPro 121
 DB 368 GCCATCAGGTTTCTGCGCAATACAGCTTTGACGGCTTGACCTTGACTGGAGTACCCTCA 427
 QY 122 GlySerArgGlySerProGlnAspLysHisLeuPheThrValLeuValGlnGluMet 141
 DB 428 GGAAGCAGGGGAGCCCTGCCGTAGACAGAGGAGCGCTTCACACCTGCTGACAGACTTG 487
 QY 142 ArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThrAla 161
 DB 488 GCCAATGCTTCCAGCAGGAAGCCAGACCTCAGGGAAGGAACGCTTCTCTGAGTGCA 547
 QY 162 AlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSerGln 181
 DB 548 GCGGTTCAGCTGGCGCAGACCTATGTGGATGCTGATACGAGGTGGACAAATGCCCGAG 607
 QY 182 TyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThr 201
 DB 608 AACCTGGATTTGTCAACCTTATGGCTTACGACTTCCATGGCTCTTGGAGAGGTCAAG 667
 QY 202 GlyGluAsnSerProLeuTyrLysTyrProThrAspThrGlySerAsnAlaTyrLeuAsn 221
 DB 668 GGACATAACAGCCCTCTCAAGAGGCAAGAGAGAGTGGTGCAGCAGCAGCCAGCTCAAC 727
 QY 222 ValAspTyrValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeuIleVal 241
 DB 728 GTGGATGCTGTGTGCAACAGTGGCTGCAGAGGGGACCCCTGCCAGCAAGCTGATCTT 787
 QY 242 GlyPheProThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIleGly 261
 DB 788 GCGATGCTTACTACGAGCGCTCTTCACTTGGCTCTCTCATCAGACACAGAGTGGGG 847
 QY 262 AlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyIleTrpAla 281
 DB 848 GCCCCAGCAGCAGGGCTGCGCACTCCAGGCCCTTCACCAAGGAAGAGGAGTGTGGCC 907
 QY 282 TyrTyrGluIleCysThrPheLeuLysAsnGlyAlaThrGlnGlyTyrAspAlaProGln 301
 DB 908 TACTATGAAGTCTGCTCTCTGG-----AAGGGGGCCACCAACAGAGAAATCCAGATCAG 961
 QY 302 GluValProTyrAlaTyrGlnGlyAsnValTrpValGlyTyrAspAsnIleLysSerPhe 321
 DB 962 AAGTGCCCTACATCTCCGGGACACAGTGGTGGTGGCTTGTATGATGTGGAGAGCTTC 1021
 QY 322 AspIleLysAlaGlnTrpLeuLysHisAsnLysPheGlyGlyAlaMetValTrpAlaIle 341
 DB 1022 AAAACCAAGGTGACGTATCTGAAGCAGAGGAGTGGGGGGGCCATGGTCTGGGCACTG 1081
 QY 342 AspLeuAspAspPheThrGlyThrPheCysAsnGlnGlyLysPheProIleIleSerThr 361
 DB 1082 GACTTAGATGACTTTGCGCGCTTCTCTGCAACAGGCGCGATACCCCTCATCCAGACG 1141
 QY 362 LeuLysLysAlaLeu-----GlyLeuGlnSerAlaSerCysThr 374
 DB 1142 CTACGCGAGGAACTGAGTCTTCCATCTTGGCTTCCAGCACCCAGAGCTTGAAGTTCCA 1201
 QY 375 AlaProAlaGlnProIleGluProIleThrAlaAlaProSerGlySerGlyAsnGlySer 394
 DB 1202 AAACCAAGTTCAGCCCTCTGAACCT----- 1225
 QY 395 GlySerSerSerSerGlyGlySerSerGlyGlySerGlyPheCysAlaValArgAlaAsn 414
 DB 1226 -----GAGCATGGGCCCCAGCCCTGGACAAGACACGTTCTGCGAGGGCAAACTGAT 1276
 QY 415 GlyLeuTyrProValAlaAsnAsnArgAsnAlaPheTrpHisCysValAsnGlyValThr 434
 DB 1277 GGGCTCTATCCCAATCTCTGGGGAACGGTCCAGCTTCTACAGCTGTGCGCGGGCGGTG 1336
 QY 435 TyrGlnGlnAsnCysGlnAlaGlyLeuValPheAspThrSerCysAspCysCysAsnTrp 454

342 AspLeuAspAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSerThr 361
 1082 GACTAGATGACTTTCGGCTTCTCTGCAACAGGCGCCATACCCCTTATCCAGACG 1141
 362 LeuLysLysAlaLeu-----GlyLeuGlnSerAlaSerCysThr 374
 1142 CTACGGCAGGAACAGTCTTCATACCTTCAGGACCCAGAGCTTGAAGTTCCA 1201
 375 AlaProAlaGlnProIleGluProIleThrAlaAlaProSerGlySerGlyAsnGlySer 394
 1202 AAACAGGTGACGCTCTGAACCT----- 1225
 395 GlySerSerSerSerGlyGlySerSerGlyGlySerGlyPheCysAlaValArgAlaAsn 414
 1226 -----GAGCATGGGCCCGCCCTGACCAAGACAGCTTCTGCAGGGCAAGCTGAT 1276
 415 GlyLeuTyProValAlaAsnAsnArgAsnAlaPheTrpHisCysValAsnGlyValThr 434
 1277 GGGCTCTATCCCAATCTCGGGAACGGTCTACAGCTGTGCAGCGGGGGGCTG 1336
 435 TyrGlnGlnAsnCysGlnAlaGlyLeuValPheAspThrSerCysAspCysCysAsnTrp 454
 1337 TTCCAGCAAGAGCTGCCGACAGGCGCTGTGTTCAGCAACTCTCTGCAAAATGCTGCACCTGG 1396

RESULT 8
 ABL57380
 ID ABL57380 standard; cDNA; 1636 BP.
 AC ABL57380;
 DT 12-AUG-2002 (first entry)
 XX Human Chitinase cDNA clone MO-218.
 DE Chitinase; enzyme; human; fungicide; antifungal; infection;
 KW candidiasis; aspergillosis; coccidioidomycosis; blastomycosis;
 KW paracoccidioidomycosis; histoplasmosis; cryptococcosis;
 KW chromoblastomycosis; sporotrichosis; mucormycosis; dermatophytosis;
 KW Pneumocystis; gene; ss.
 OS Homo sapiens.

Key Location/Qualifiers
 CDS 2..1402
 /tag= a
 /product= "Chitinase"
 sig_peptide 2..64
 /tag= b
 mat_peptide 65..1399
 /tag= c
 US6372212-B1.
 XX 16-APR-2002.
 XX 16-JUN-1997; 97US-0877599.
 XX 14-JUN-1996; 96US-0663618.
 XX (ICOS-) ICOS CORP.
 XX Gray PW;
 XX WPI; 2002-442449/47.
 XX P-PSDB; ABB76291.
 XX Co-administering chitinase to improve the effectiveness of fungicidal
 XX drugs e.g. amphotericin B or itraconazole, useful for treating fungal
 XX infections e.g. candidiasis, coccidioidomycosis and blastomycosis -
 XX Example 1; Column 23-28; 26pp; English.

The present sequence is that of cDNA clone MO-218 (ATCC 98077) encoding human chitinase (see ABB76291). The clone was isolated from a cDNA library prepared from peripheral blood monocyte-derived macrophages following sequence analysis. Also isolated was clone MO-13B (see ABL57381), which contains a single nucleotide difference in the coding region, changing the encoded amino acid at position 81 of the mature protein from glycine to serine. Northern blots showed highest chitinase gene expression in lung and ovary tissues. Expression in lung is consistent with a protective role against pathogenic organisms that contain chitin. The invention provides human chitinase polynucleotides and polypeptides, and materials and methods for the recombinant production of human chitinase products, which are expected to be useful as products for treating fungal infections or for the development of such products. Human chitinase has a synergistic effect on the actions of other fungicides. It can be administered to improve the antifungal activity of a non-chitinase antifungal agent, especially amphotericin B or itraconazole, in the treatment of a fungal infection such as candidiasis, aspergillosis, coccidioidomycosis, blastomycosis, paracoccidioidomycosis, histoplasmosis, cryptococcosis, chromoblastomycosis, sporotrichosis, mucormycosis, dermatophytoses and Pneumocystis infections (all claimed). In particular, the fungal infection involves Candida, Aspergillus and/or Cryptococcus spp., whose growth is not effectively inhibited by contact with human chitinase alone.

XX
 SQ Sequence 1636 BP; 361 A; 491 C; 440 G; 344 T; 0 other;

Alignment Scores:
 Pred. No.: 2,35e-111 Length: 1636
 Score: 1323.00 Matches: 240
 Percent Similarity: 67.39% Conservative: 70
 Best Local Similarity: 52.17% Mismatches: 126
 Query Match: 52.79% Indels: 24
 DB: 24 Gaps: 3

US-10-004-219B-14 (1-455) x ABL57380 (1-1636)

QY 2 GlnLeuThrCysTyThrPheThrAsnTrpAlaGlnTyArgProGlyLeuGlyArgPheMet 21
 DB 68 AAACCTGGTCTGCTACTTCCCACTGGGCCCCAGTACACAGAGGGGAGGCTCGCTTCCTG 127

QY 22 ProAspAsnIleAspProCysLeuCysThrHisLeuIleTyAlaPheAlaGlyArgGln 41
 DB 128 CCCAAGGACTTGGACCCCGAGCTTTGCACCCACCTCATCTAGCCCTTCGCTGGCATGCC 187

QY 42 AsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyThrGlnAlaPheAsnGly 61
 DB 188 AACCCACAGCTGAGCACCACCTGAGTGGAAATGACGAGACTCTCTACAGGAGTTCAATGGC 247

QY 62 LeuLysAsnLysAsnSerGlnLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPheGly 81
 DB 248 CTGAAGAAGATGAATCCCAAGCTGAAGACCTGTGTAGCCATCGGAGGCTGGAAATTCGGC 307

QY 82 ThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThrSer 101
 DB 308 ACTCAGAAGTTTCACAGATATGCTAGCCACGCGCAACACCGTCACACCTTTGTCAACTCG 367

QY 102 ValIleLysPheLeuArgGlnTyThrGluPheAspGlyLeuAspPheAspTrpGluTyPro 121
 DB 368 GCCATCAGGTTTCTGCGCAATATACAGCTTTGACGCGCTTGACCTTGACCTGGGAGTACCCA 427

QY 122 GlySerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValGlnGluMet 141
 DB 428 GGAAGCCAGGGAGGCCCTTCGCTAGCAAGAGGCGCTTCACACCCCTGGTACGAGACTTG 487

QY 142 ArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThrAla 161
 DB 488 GCCAATGCTTCCACAGCAGAGAGCCAGACCTCAGGGAAGGAACGCTTCTCTGAGTGCA 547

QY 162 AlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyThrGluIleProGlnLeuSerGln 181
 DB 548 GCGGTTCCAGCTGGGCGAGACCTATGTGGATGCTGTGATACGAGGTGACAAATTCGCCAG 607


```
QY 102 ValIleLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyrPro 121
DB 379 GCCATCAGGTTCTGCGCAATACAGCTTTGACGGCCTTGACTGGAGTACCCA 438
QY 122 GlySerArgGlySerProGlnAspLysHisLeuPheThrValLeuValGlnGluMet 141
DB 439 GGAAGCCGAGGAGCCCTCCCTAGCAAGAGCGCTTCACAAACCTGGTACAGGACTTG 498
QY 142 ArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThrAla 161
DB 499 GCCATGCTTTCAGCAGAGAGCCAGACTCAGGAAGAGACGCTTCTTCTGAGTGA 558
QY 162 AlaValAlaAlaGlyLysSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSerGln 181
DB 559 GCGGTTCACGCTGGGCAGACCTATGTGGATGCTGGATACGAGTGGACAAAATCGCCAG 618
QY 182 TyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThr 201
DB 619 AACCTGGATTTTGTCAACCTTATGCGCTACGACTTCCATGGCTCTTGGGAGAGTCA 678
QY 202 GlyGluAsnSerProLeuTyrLysTyrProThrAspThrGlySerAsnAlaTyrLeuAsn 221
DB 679 GGACATACAGCCCTCTACAGAGCGAAGAGAGTGGTGCAGCAGCCAGCCTCAAC 738
QY 222 ValAspTyrValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeuIleVal 241
DB 739 GTGGATGCTGCTGTGCAACAGTGGCTGCAGAGGGGAGCCCTGCCAGCAAGCTGATCTT 798
QY 242 GlyPheProThrTyrGlyHisAsnPhelLeuSerAsnProSerAsnThrGlyIleGly 261
DB 799 GGCATGCCCTACCTACGGAGCGCTCTTCACACTGGCCTCTCATCAGACACACAGAGTGG 858
QY 262 AlaProThrSerGlyValGlyProAlaGlyProTyrAlaLysGluSerGlyIleTrpAla 281
DB 859 GCCCAGCCAGGAGGTCTGGCACTCAGGCCCTTCCACCAAGAGAGGAGGATGCTGGCC 918
QY 282 TyrTyrGluIleCysThrPheLeuLysAsnGlyAlaThrGlnGlyTyrAspAlaProGln 301
DB 919 TACTATAGTCTGCTCTGCG-----AAGGGGGCCACCAACAGAGATCCAGGATCAG 972
QY 302 GluValProTyrAlaTyrGlnGlyAsnValTrpValGlyTyrAspAsnIleLysSerPhe 321
DB 973 AAGGTGCCCTACATCTTCGGGACCAACCACTGGTGGTGGCTTTGATGATGTGGAGACTTC 1032
QY 322 AspileLysAlaGluTrpLeuLysHisAsnLysPheGlyValAlaMetValTrpAlaIle 341
DB 1033 AAACCAAGGTGAGCTATCTGAAGCAGAAAGGAGCTGGGGGGCCATGCTCTGGGCACATG 1092
QY 342 AspLeuAspAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSerThr 361
DB 1093 GACTTAGATGACTTTGCGGCTTCTCTGCAACAGGCGCGATACCCCTCATCCAGAGC 1152
QY 362 LeuLysLysAlaLeu-----GlyLeuGlnSerAlaSerCysThr 374
DB 1153 CTAGCGAGGAACCTAGTCTTCATACTTGCCTTCCAGGCACCCAGAGCTTCAAGTTCCA 1212
QY 375 AlaProAlaGlnProIleGluProIleThrAlaAlaProSerGlySerGlyAsnGlySer 394
DB 1213 AAACCAAGGTGACCCCTCTGAACCT----- 1236
QY 395 GlySerSerSerSerGlySerSerGlySerGlyPheCysAlaValArgAlaAsn 414
DB 1237 -----GAGCATGGCCCGCCAGCTTGCAGACACAGCTTCTGCAGGCGCAAGCTGAT 1287
QY 415 GlyLeuTyrProValAlaAsnAsnArgAsnAlaPheTrpHisCysValAsnGlyValThr 434
DB 1288 GGGCTCTATCCCAATCTCTCGGAAGCGTCCAGCTTCTACAGCTGTGCAGCGGGCGGCTG 1347
QY 435 TyrGlnGlnAsnCysGlnAlaGlyLeuValPheAspThrSerCysAspCysAsnTrp 454
DB 1348 TTCCAGCAAGACTGCCCGCAGCGCTGGTGTTCAGCAACTCTCTGCAAAATGCTGCACCTGG 1407
```

```
RESULT 10
AAV10436
ID AAV10436 standard; cDNA; 1656 BP.
XX
AC AAV10436;
XX
DT 15-JUN-1998 (first entry)
XX
DE Human chitinase clone MO-13B cDNA.
XX
KW Chitinase; human; fungal infection; immunogen; diagnosis; treatment;
KW Gaucher's disease; transgenic; detection; hybridisation; antifungal;
KW Rheumatoid arthritis; overexpression; extracellular matrix; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 27..1427
FT /*tag= a
FT sig_peptide 27..89
FT /*tag= b
FT mat_peptide 90..1424
FT /*tag= c
FT /product= chitinase
FT /note= "from clone MO-13B"
XX
XX WO9747752-A1.
XX
XX 18-DEC-1997.
XX
XX 16-JUN-1997; 97WO-US10460.
XX
XX 14-JUN-1996; 96US-0663618.
XX
XX (ICOS-) ICOS CORP.
XX
XX Gray PW;
XX
XX WPI; 1998-052316/05.
XX
XX P-PSDB; AAW40260.
XX
XX Nucleic acids encoding human chitinase - useful as antifungal
XX agents, especially in combination with other antifungals
XX
XX Claim 9; Page 42-44; 63pp; English.
XX
XX This sequence encodes a novel human chitinase isolated from clone MO-13B.
XX Chitinases are useful for treating or preventing fungal infection and
XX as immunogens for generating antibodies which are used to purify, detect
XX and quantify chitinases, e.g. for diagnosis of Gaucher's disease. The
XX nucleic acid sequence of the chitinase is also useful as a probe to
XX identify and isolate genomic DNA encoding chitinases or similar proteins,
XX or cells expressing them or to generate transgenic ('knockout') rodents.
XX It can also be used in hybridisation assays and to detect genetic
XX alterations in the chitinase gene related to disease. Agents that
XX this protein may be useful in treatment of Gaucher's disease and
XX rheumatoid arthritis, where overexpression of the protein can damage
XX the extracellular matrix. Chitinase also improves the activity of other
XX antifungal agents and may allow a reduction in the dose of such agents,
XX and thus of their side effects.
XX
XX Sequence 1656 BP; 365 A; 497 C; 447 G; 347 T; 0 other;
XX
Alignment Scores:
Score: 8.46e-111 Length: 1656
Pred. No.: 1317.00 Matches: 239
Percent Similarity: 67.17% Conservative: 70
Best Local Similarity: 51.96% Mismatches: 127
Query Match: 52.55% Indels: 24
DB: 19 Gaps: 3
XX
US-10-004-219B-14 (1-455) x AAV10436 (1-1656)
```

QY 2 GlnLeuThrCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPheMet 21
DB : : : : :
93 AAACGTGCTGCTACTTCAACCACTGGGCCAGTACAGACAGGGAGGCTGCTTCCTG 152
QY 22 ProAsnLeuLeuAspProCysLeuCysThrHisLeuLeuLeuTyrAlaPheAlaGlyArgGln 41
DB : : : : :
153 CCAAGAGACTTGGACCCAGCCCTTTCACCCACTCATCTACGCCCTTCGCTGGCATGACC 212
QY 42 AsnAsnGluLeuThrThrLeuGluTrpAsnAspValTrpLeuTyrGlnAlaPheAsnGly 61
DB : : : : :
213 AACACAGCTGAGCACCACTGAGTGGATGACAGACTCTCTACAGAGATTCATATGC 272
QY 62 LeuLysAsnLysAsnSerGlnLeuLysThrLeuLeuAlaLeuLeuGlyTrpAsnPheGly 81
DB : : : : :
273 CTGAAGAAGATGAATCCCAAGCTGAAGACCCCTGTAGCCATCGAGGCTGAATTTTCAGC 332
QY 82 ThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheLeuThrSer 101
DB : : : : :
333 ACTCAGAAGTTACAGATATGGTAGCCACGCCCAACACCGTCAGACCTTTGTCAACTCG 392
QY 102 ValLeuLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyrPro 121
DB : : : : :
393 GCCATCAGTTCTTGGCAATATACAGCTTGAAGGCTTGACCTGACTGGAGTACCCA 452
QY 122 GlySerArgGlySerProGlnAspLysHisLeuPheThrValLeuValGlnGluMet 141
DB : : : : :
453 GGAAGCCAGGGAGCCCTGCGGTAGACAGAGCGCTTCACAACCCCTGGTACAGGACTTG 512
QY 142 ArgGluAlaPheGluGlnGluAlaLysGlnLeuAsnLysProArgLeuMetValThrAla 161
DB : : : : :
513 GCCAATGCCCTTCAGCAGGAAGCCCAAGCCTCAGGGAAGGAGCGCTTCTTCTGAGTGA 572
QY 162 AlaValAlaAlaGlyLeuSerAsnLeuGlnSerGlyTyrGluLeuProGlnLeuSerGln 181
DB : : : : :
573 GCGGTTCCAGCTGGCAGACCTATGTGATCTGGATACAGGTGACAAATCGCCAG 632
QY 182 TyrLeuAspTyrLeuHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThr 201
DB : : : : :
633 AACCTGGATTTGTCAACCTTATGGCTACGACTTCATGGCTCTTGGGAGAGGTTCAG 692
QY 202 GlyLeuAsnSerProLeuTyrLysTyrProThrAspThrGlySerAsnAlaTyrLeuAsn 221
DB : : : : :
693 GGACATAACAGCCCTCTACAGAGCGCAAGAGAGAGTGGTCGACGACGACCGCTCAAC 752
QY 222 ValAspTyrValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeuLeuVal 241
DB : : : : :
753 GTGATGCTGCTGCAACAGTGGCTGCAGAGGGGACCCCTGCCAGACGATCCTT 812
QY 242 GlyPheProThrTyrGlyHisAsnPheLeuSerAsnProSerAsnThrGlyLeuGly 261
DB : : : : :
813 GGCATGCTACCTACGACGCTCCTTCACACTGGCTCCTCTCATCAGACACAGAGTGGG 872
QY 262 AlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyLeuTrpAla 281
DB : : : : :
873 GCCCCAGCACAGGGTCTGCACTCCAGGCCCTCTCACAAGGAAGAGGAGTGGTGGCC 932
QY 282 TyrTyrGluLeuCysThrPheLeuLysAsnGlyAlaThrGlnGlyTrpAspAlaProGln 301
DB : : : : :
933 TACTATGAGTCTGCTCTGG-----AAGGGGGCCCAACCAACAGAGATCCAGGATCAG 986
QY 302 GluValProGlyTyrAlaTyrGlnGlyAsnValTrpValGlyTyrAspAsnLeuLysSerPhe 321
DB : : : : :
987 AAGTGCCCTTACATCTTCGGGCAACACCACTGAGTGGTGGCTTTGATGATGTGGAGACTTC 1046
QY 322 AspLeuLeuAlaGlnTrpLeuLysHisAsnLysPheGlyGlyAlaMetValTrpAlaLeu 341
DB : : : : :
1047 AAACCAAGTCACTATCTGAAGCAGAGAGGAGCTGGCGGGGCCAATGCTGTGGCACTG 1106
QY 342 AspLeuAspAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuLeuSerThr 361
DB : : : : :
1107 GACTTAGATGACTTTCGGGCTCTCTCTGCAACAGGCGCGATACCCCTCATCCAGAGC 1166
QY 362 LeuLysLysAlaLeu-----GlyLeuGlnSerAlaSerCysThr 374

DB 1167 CTACGGCAGGAACCTGAGTCTTCATCTGCTTCAGGCACCCAGAGCTTGAAGTTCCA 1226
QY 375 AlaProAlaGlnProIleGluProIleThrAlaAlaProSerGlySerGlyAsnGlySer 394
DB 1227 AAACGAGTGCAGCCCTCTGAACCT----- 1250
QY 395 GlySerSerSerGlyGlySerSerGlyGlySerGlyPheCysAlaValArgAlaAsn 414
DB : : : : :
1251 -----GAGCATGGCCCCAGCCCTGGACAAGACACGTTCTGCCAGGGCAAGCTGAT 1301
QY 415 GlyLeuTyrProValAlaAsnAsnArgAsnAlaPheTrpHisCysValAsnGlyValThr 434
DB : : : : :
1302 GGGCTCTATCCCAATCTCGGGAACGGTCCAGCTTCTACAGCTGTGCAGCGGGGGCTG 1361
QY 435 TyrGlnGlnAsnCysGlnAlaGlyLeuValPheAspThrSerCysAspCysCysAsnTrp 454
DB : : : : :
1362 TTCCAGCAAGAGCTGCCGACAGGCTGGTGTTCAGCAACTCTGCAAAATGCTGCACCTGG 1421
RESULT 11
AAZ21848
ID AAZ21848 standard; DNA; 1656 BP.
XX
AC AAZ21848;
XX
XX 10-DEC-1999. (first entry)
DT
XX
DE MO-13B clone of human Chitinase, with noncoding 5'/3' regions.
XX
XX chitin; fungal infection; immunocompromised; AIDS; chemotherapy;
KW organ transplant; parasite; chitin-binding; allele; vector;
KW truncated protein; ds.
OS
XX Homo sapiens.
FH Key Location/Qualifiers
FT CDS 27..1427
FT /tag= a
FT /product= Human_Chitinase
FT sig_peptide 27..89
FT /tag= b
FT /note= "Signal peptide"
FT mat_peptide 90..1427
FT /tag= c
FT /note= "Mature peptide"
XX
XX WO9946390-A1.
PN
XX
PD 16-SEP-1999.
XX
XX 12-MAR-1999; 99WO-US05343.
PF
XX
PR 12-MAR-1998; 98US-0039198.
XX
XX (ICOS-) ICOS CORP.
PA
XX
XX Gray PW, Tjoelker LW;
XX
XX WPI; 1999-551417/46.
DR P-PSDB; AAY42426.
XX
XX Novel chitin-binding fragments of human chitinase used to treat fungal
PT infections in animals
XX
XX Example 1; Page 59-62; 83pp; English.
PS
XX
XX This is the nucleotide sequence of an allelic form of the human
CC chitinase enzyme, which is capable of degrading Chitin (a linear
CC homo polymer of beta-1,4-linked N-acetylglucosamine residues).
CC Chitinase fragments can be used to screen for proteins or other
CC molecules that specifically bind to the chitin-binding domain of human
CC chitinase or that modulate its activity. These compounds are useful for
CC immunization, as well as for purifying chitinase, as well as for

Qy 222 ValAspTyrvalMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeuIleVal 241

FT XX /product= "Human mature chitinase from clone pMO-13B"
PN XX WO200123430-A2.
XX XX
PD 05-APR-2001.
XX 28-SEP-2000; 2000WO-US26960.
XX 30-SEP-1999; 99US-0409918.
PR (ICOS-) ICOS CORP.
XX
XX Allison DS, Dietsch GN, Gray PW, Shaw KD, Steiner BH;
PI WPI: 2001-266141/27.
DR P-PSDB; ABE00433.
DR
XX Novel chitinase immunoglobulin fusion product, useful for treating
PT fungal infections and reducing the amount of a non-chitinase antifungal
PT agent needed for the treatment
XX
XX Claim 2; Page 34-36; 39pp; English.
XX
XX The present invention relates to a chitinase immunoglobulin (Ig) fusion
CC product, comprising a human chitinase fused to at least a portion of an
CC immunoglobulin chain. The fusion product is useful for treating fungal
CC infections (mycoses) such as candidiasis, aspergillosis, blastomycosis,
CC coccidioidomycosis, paracoccidioidomycosis, histoplasmosis, mucormycosis,
CC cryptococcosis, chromoblastomycosis, sporotrichosis and dermatophytoses.
CC The fusion protein is useful for reducing the amount of non-chitinase
CC antifungal agent needed to exert an antifungal activity. The fusion
CC protein is also useful for preparing a medicament for the prophylactic
CC or therapeutic treatment of fungal infections. Chitinase immunoglobulin
CC fusion product has unexpectedly improved serum half-life and formulation
CC properties. The present sequence is human chitinase cDNA from clone
CC pMO-13B. Chitinase enzyme degrades chitin which is a homopolymer of
CC beta-(1,4)-linked N-acetylglucosamine residues.
XX
SQ Sequence 1656 BP; 365 A; 497 C; 447 G; 347 T; 0 other;

Alignment Scores:
Pred. No.: 8,46e-111 Length: 1656
Score: 1317.00 Matches: 239
Percent Similarity: 67.17% Conservative: 70
Best Local Similarity: 51.96% Mismatches: 127
Query Match: 52.55% Indels: 24
DB: 22 Gaps: 3
US-10-004-219B-14 (1-455) x AAD03760 (1-1656)
QY 2 GlnLeuThrCysThrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPheMet 21
Db 93 AACTGGTCTGCTACTTACCACTAGGCGCCAGTACAGACAGGGGAGGCTCGCTTCCTG 152
QY 22 ProAspAsnIleAspProCysLeuGlyCysThrHisLeuIleTyrAlaPheAlaGlyArgGln 41
Db 153 CCCAAGGACTTGGACCCAGCCTTTGCACCCACCTCATCTAGCGCTTGGCTGGCATGACC 212
QY 42 AsnAsnGluLeuThrThrIleGluTrpAsnAspValThrLeuTyrGlnAlaPheAsnGly 61
Db 213 AACCCAGCCTGAGCACCCTAGTGGTGAATGACGAGACTCTCTACAGGAGTTCAATGGC 272
QY 62 LeuLeuAsnLysAsnSerGlnLeuLysThrLeuLeuAlaIleGlyGlyTyrAsnPheGly 81
Db 273 CTGAGAGAGTGAATCCAGCTGAGAGCCCTTTAGCCATCCGAGGCTGGAATTTTCAGC 332
QY 82 ThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThrSer 101
Db 333 ACTCAGAAGTTTACAGATATGGTAGCCAGCCGCAACCGTCAGACCTTTGTCAACTCG 392
QY 102 ValIleLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyrPro 121
Db 393 GCCATCAGGTTTCTGGCGCAATACAGCTTTGACGGGCTTGACCTTGACTGGGAGTACCCA 452

QY 122 GlySerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValGlnGluMet 141
Db 453 GGAAGCCAGGGGAGCCCTGCGGTAGACAAGGAGCGCTTACAAACCTGGTGTGACGACTTG 512
QY 142 ArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThrAla 161
Db 513 GCCAATGCGCTTCCAGCAGGAGCCAGCCTCAGGAGGAGGAGGAGGAGGAGGAGGAGG 572
QY 162 AlaValAlaAlaGlyLeuSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSerGln 181
Db 573 GCGGTTCCAGCTGGGCGACACCTATGTGGATGCTGGATACGAGGTGACAAATGCGCCAG 632
QY 182 TyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThr 201
Db 633 AACCTGGATTTTGTCAACCTTATGCGCTTACGACTTCCATGGCTCTTGGGAGAAGGTTCAG 692
QY 202 GlyGluAsnSerProLeuTyrIleGlyTyrProThrAspThrGlySerAsnAlaTyrLeuAsn 221
Db 693 GGACATAACAGCCCGCTCTACAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 752
QY 222 ValAspTyrValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeuIleVal 241
Db 753 GTGGATGCTGCTGTGCAACAGTGGCTGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 812
QY 242 GlyPheProThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIleGly 261
Db 813 GGCATGCGCTTACCTACGAGCGCTTCTTACACTGGCTCTCATCAGACACCAAGAGTGGG 872
QY 262 AlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyIleTrpAla 281
Db 873 GCGCCAGCAGCAGGCTTGGCATCCTCAGCGCCCTTCCACCAAGGAGGAGGAGGAGGAG 932
QY 282 TyrTyrGluIleCysThrPheLeuLysAsnGlyAlaThrGlnGlyTyrAspAlaProGln 301
Db 933 TACTATGAAGTCTGCTCTCTGG-----AAGGGGGGCCACCAACACAGAGATCCAGATCAG 986
QY 302 GluValProTyrAlaTyrGlnGlyAsnValTrpValGlyTyrAspAsnIleLysSerPhe 321
Db 987 AAGGTGCCCTACATCTCCGGGACAAACCACTGGTGGGCTTGTGATGATGTGGAGAGCTTC 1046
QY 322 AsnIleLysAlaGlnTrpLeuLysHisAsnLysPheGlyGlyAlaMetValTrpAlaIle 341
Db 1047 AAAACCAAGGTCAGCTATCTGAAGCAGAGAGGAGCTGGGCGGGGCCATGCTCTGGGCACTG 1106
QY 342 AspLeuAspAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSerThr 361
Db 1107 GACTTAGATGACTTTGGCGGCTTCTCTGCAACCAAGGCGCGGATACCCCTCATCCAGACG 1166
QY 362 LeuLysLysAlaLeu-----GlyLeuGlnSerAlaSerCysThr 374
Db 1167 CTACGGCAGGAAGTCTGCTTCCATCTTCCATCTTCCAGGCAACCCAGAGCTTGAAGTTCCA 1226
QY 375 AlaProAlaGlnProIleGluProIleThrAlaAlaProSerGlySerGlyAsnGlySer 394
Db 1227 AAACCAAGGTCAGCCCTCTGAACCT----- 1250
QY 395 GlySerSerSerSerGlyGlySerSerGlyGlySerGlyPheCysAlaValArgAlaAsn 414
Db 1251 -----GACATGGCGCCCGCCCTGGACAGACAGCTTCTGCCAGGGCAAGCTGAT 1301
QY 415 GlyLeuTyrProValAlaAsnAsnArgAsnAlaPheTrpHisCysValAsnGlyValThr 434
Db 1302 GGGCTCTATCCCAATCTCTGGGAACGGTCCAGCTTCTACAGCTGTGCAGCGGGCGGCTG 1361
QY 435 TyrGlnGlnAsnCysGlnAlaGlyLeuValPheAspThrSerCysAspCysCysAsnTrp 454
Db 1362 TTCCAGCAAGGCTGGCCCGCAGCAGGCTGTGTTCAGCAACTCTCTGCAATGCTGCACCTG 1421
RESULT 13
ABL57381
ID ABL57381 standard; cDNA; 1656 BP.
XX

ABL57381;
 12-AUG-2002 (first entry)
 Human chitinase cDNA clone MO-13B.
 Chitinase; enzyme; human; fungicide; antifungal; infection;
 candidiasis; aspergillosis; coccidioidomycosis; blastomycosis;
 paracoccidioidomycosis; histoplasmosis; cryptococcosis;
 chromoblastomycosis; sporotrichosis; mucormycosis; dermatophytosis;
 Pneumocystis; gene; ss.
 Homo sapiens.
 Key Location/Qualifiers
 CDS 27..1427 /tag= a
 /product= "Chitinase"
 sig_peptide 27..89 /tag= b
 mat_peptide 90..1424 /tag= c
 US6372212-B1.
 16-APR-2002.
 16-JUN-1997; 97US-0877599.
 14-JUN-1996; 96US-0663618.
 (ICOS-) ICOS CORP.
 Gray PW;
 WPI; 2002-442449/47.
 P-PSDB; ABB76292.
 Co-administering chitinase to improve the effectiveness of fungicidal
 drugs e.g. amphotericin B or itraconazole, useful for treating fungal
 infections e.g. candidiasis, coccidioidomycosis and blastomycosis -
 Example 1; Column 31-34; 26pp; English.
 The present sequence is that of cDNA clone MO-13B (ATCC 98078)
 encoding human chitinase (see ABB76292). The clone was isolated
 from a cDNA library prepared from peripheral blood monocyte-derived
 macrophages using a probe based on human chitinase clone MO-218
 (see ABL57380). The 2 clones contain a single nucleotide
 difference in the coding region, causing amino acid position 81 of
 the mature protein to be either serine or glycine. Northern blots
 showed highest chitinase gene expression in lung and ovary tissues.
 Expression in lung is consistent with a protective role against
 pathogenic organisms that contain chitin. The invention provides
 human chitinase polynucleotides and polypeptides, and materials and
 methods for the recombinant production of human chitinase products,
 which are expected to be useful as products for treating fungal
 infections or for the development of such products. Human
 chitinase has a synergistic effect on the actions of other
 fungicides. It can be administered to improve the antifungal
 activity of a non-chitinase antifungal agent, especially
 amphotericin B or itraconazole, in the treatment of a fungal
 infection such as candidiasis, aspergillosis, coccidioidomycosis,
 blastomycosis, paracoccidioidomycosis, histoplasmosis,
 cryptococcosis, chromoblastomycosis, sporotrichosis, mucormycosis,
 dermatophytosis and Pneumocystis infections (all claimed). In
 particular, the fungal infection involves Candida, Aspergillus
 and/or Cryptococcus spp., whose growth is not effectively
 inhibited by contact with human chitinase alone.
 Sequence 1656 BP; 365 A; 497 C; 447 G; 347 T; 0 other;
 Alignment Scores:

Pred. No.:	8.46e-111	Length:	1656
Score:	1317.00	Matches:	239
Percent Similarity:	67.17%	Conservative:	70
Best Local Similarity:	51.96%	Mismatches:	127
Query Match:	52.55%	Indels:	24
DB:	24	Gaps:	3

US-10-004-219b-14 (1-455) x ABL57381 (1-1656)

QY	2	GlnLeuThrCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPheMet	21
DB	93	AAACTGGTCTGCTACTTACCACACTGGGCGCCAGTACACAGAGGGAGGCTCGCTTCCTG	152
QY	22	ProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArgGln	41
DB	153	CCCAAGGACTTGGACCCCGAGCTTTGCACCCACCTCATCTAGCCCTTCGTGGCATGACC	212
QY	42	AsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrGlnAlaPheAsnGly	61
DB	213	AACCAACAGCTGAGCACCACCTGAGTGGAAATGACGAGACTCTTACCAAGAGTTCAATGC	272
QY	62	LeuLysAsnLysAsnSerGlnLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPheGly	81
DB	273	CTGAAGAAAGATGAATCCCAAGCTGAAGACCTGTGTAGCCATCGGAGGCTGGAAATTCAGC	332
QY	82	ThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThrSer	101
DB	333	ACTCAGAAGTTTCACAGATATGGTAGCCACGGCCCAACACCGTCAGACCTTTGTCAACTCG	392
QY	102	ValIleLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyrPro	121
DB	393	GCCATCAGGTTTCTGCGCAATACAGCTTTGACGGCCTTGACCTTGACTGGGAGTACCCA	452
QY	122	GlySerArgGlySerProGlnAspLysHisLeuPheThrValLeuValGlnGluMet	141
DB	453	GGAAGCCAGGGGAGCCCTGCGGTAGACAAGAGGCGCTTCACAACCCCTGGTACAGACTTG	512
QY	142	ArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThrAla	161
DB	513	GCCAAATGCTTCCAGCAGAGAGCCAGACCTCAGGAGAGGAACGGCTTCTTCTGAGTGCA	572
QY	162	AlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSerGln	181
DB	573	GCGGTTCAGCTGGCGCAGACCTATGTGATGTGATGATGATGATGATGATGATGATGATG	632
QY	182	TyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluClyTyrThr	201
DB	633	AACCTGGATTTTGTCAACCTTATGGCCCTACGACTTCCATGGCTCTTGGGAGAGTCACG	692
QY	202	GlyGluAsnSerProLeuTyrLysTyrProThrAspThrGlySerAsnAlaTyrLeuAsn	221
DB	693	GGACATACAGCCCTCTTACAGAGGCAAGAGAGAGTGGTGCAGCAGCCAGCTCAAC	752
QY	222	ValAspTyrValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeuVal	241
DB	753	GTGGATGTGCTGTGCAACCTGGTGTGCAAGAGGCGACCCCTGCCAGCAAGTGTATCTT	812
QY	242	GlyPheProThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrClyIleGly	261
DB	813	GGCATGCCCTACCTAGCGACGCTCTTCACTGGCCCTCATCAGACCAAGAGTGGGG	872
QY	262	AlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyIleTrpAla	281
DB	873	GCCCAAGCCAGGAGGTCTGGACCTCCAGGCCCTTCCACAGAGAGAGGAGGTGTGGCC	932
QY	282	TyrTyrGluIleCysThrPheLeuLysAsnGlyAlaThrGlnGlyTrpAspAlaProGln	301
DB	933	TACTATGAAGTCTGCTCTCTGG-----AAGGGGGCCACCAACACAGAGAATCCAGATCAG	986
QY	302	GluValProTyrAlaTyrGlnGlyAsnValTrpValGlyTyrAspAsnIleLysSerPhe	321
DB	987	AAGGTGCCCTACATCTCTCCGGGACAAACAGTGGTGGGCTTTTGATGATGTGGAGACTTC	1046

Alignment Scores:

Db 910 GGCATGCGCTACCTACGAGCGCTCTTCCACATCGGCTCTCATCAGACACAGAGTGGG 969
 Qy 262 AlaProThrSerGlyValGlyProAlaGlyProTyrAlaValGlySerGlyIleTrpAla 281
 Db 970 GCGCCAGCCAGGGTCTGGCACTCAGGCCCTTCCACCAAGGAAGGAGGATGCTGGCC 1029
 Qy 282 TyrTyrGluIleCysThrPheLeuLysAsnGlyAlaThrGlnGlyTrpAspAlaProGln 301
 Db 1030 TACTATGAAGTCTGCTCTGG-----AAGGGGGCCACCAACAGAGATCCAGGATCAG 1083
 Qy 302 GluValProTyrAlaTyrGlnGlyAsnValTrpValGlyTyrAspAsnIleLysSerPhe 321
 Db 1084 AAGGTGCCCTACATCTTCGGGACCAACCAAGTGGTGGCTTTGATCATGTGGAGAGCTTC 1143
 Qy 322 AspIleLysAlaGlnTrpLeuLysHisAsnLysPheGlyGlyValMetValTrpAlaIle 341
 Db 1144 AAACCAAGGTGAGCTATCTGAAGCAAGAGGACTGGGGCGGCCATGCTCTGGGACTG 1203
 Qy 342 AspLeuAspAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSerThr 361
 Db 1204 GACTTAGATGACTTTGGCGGCTTCTCTGCAACCAAGGCGCGATACCCCTCATCCAGAGC 1263
 Qy 362 LeuLysLysAlaLeu-----GlyLeuGlnSerAlaSerCysThr 374
 Db 1264 CTACGGCAGGAAGTCTGCTTCCATCTTCCATCTTCCAGGCACCCAGAGCTTGAAGTTCCA 1323
 Qy 375 AlaProAlaGlnProIleGluProIleThrAlaAlaProSerGlySerGlyAsnGlySer 394
 Db 1324 AAACCAAGTCAAGCTCTGAAGCT-----1347
 Qy 395 GlySerSerSerGlyGlySerGlySerGlySerGlyPheCysAlaValArgAlaAsn 414
 Db 1348 -----GAGCATGGCCCGCCCGCTGCAAGACACGCTTCTGCCAGGGCAAGCTGAT 1398
 Qy 415 GlyLeuTyrProValAlaAsnAsnArgAsnAlaPheTrpHisCysValAsnGlyValThr 434
 Db 1399 GGGCTCTATCCCAATCTCTGGGAACGGTCTCAGCTTCTACAGCTGTGACGCGGGGGCTG 1458
 Qy 435 TyrGlnGlnAsnCysGlnAlaGlyLeuValPheAspThrSerCysAspCysCysAsnTrp 454
 Db 1459 TTCCAGCAAGCTGCCCGACAGCGCTGGTGTTCAGCAACTCTCTGCAAAATGCTGCACCTGG 1518
 RESULT 15
 AAT50834
 ID AAT50834 standard; cDNA; 1713 BP.
 XX
 AC AAT50834;
 XX
 DT 24-MAR-1997 (first entry)
 XX
 DE Human chitinase cDNA clone chi.39.
 XX
 KW Chitinase; chitotriosidase; chitin; infectious disease;
 KW gene therapy; vaccine; lysosomal lipidoses; Gaucher's disease;
 KW leishmaniasis; sarcoidosis; X-linked adrenoleukodystrophy;
 KW multiple sclerosis; drug delivery; cosmetics; food; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 13..1176
 FT /*tag= a
 FT sig_peptide 13..75
 FT /*tag= b
 FT mat_peptide 76..1173
 FT /*tag= c
 XX
 PN WO9640940-A2.
 XX
 PD 19-DEC-1996.
 XX
 PF 06-JUN-1996; 96WO-NL00225.

XX 07-JUN-1995; 95US-0486839.
 PR (UNAM) UNIV AMSTERDAM.
 XX
 PA Aerts JMG;
 XX
 PI WPI; 1997-118698/11.
 DR P-PSDB; AAW08585.
 XX
 DR New human chitinase - used to treat or prevent infection by
 PT chitin-contg. pathogens, in diagnosis and as additives to cosmetics,
 PT foods, implants etc.
 XX
 PS Claim 2; Page 42-43; 58pp; English.
 CC A cDNA clone (AAT50834), designated chi.39, codes for a 39 kDa human
 CC chitinase (AAW08584) that is stable to many proteases, active at pH 3-
 CC 8 and up to 50 deg, and stable in the circulation. Clones chi.39
 CC and chi.50 (see also AAT50833) were isolated from a human macrophage
 CC cDNA library using as probe a partial clone obtd. using primers
 CC (see also AAT50835-36) based on a chitotriosidase purified from a
 CC type 1 Gaucher disease patient. The 2 cDNA clones are the result
 CC of alternative splicing of RNA. Chitinase nucleic acid can be used
 CC for large-scale prodn. of recombinant human chitinases, or can be
 CC incorporated into a gene therapy vector to treat or prevent
 CC infection by chitin-contg. pathogens.
 XX
 SQ Sequence 1713 BP; 380 A; 503 C; 465 G; 365 T; 0 other;
 Alignment Scores:
 Pred. No.: 2,82e-110 Length: 1713
 Score: 1311.50 Matches: 244
 Percent Similarity: 68.03% Conservative: 76
 Best Local Similarity: 51.91% Mismatches: 131
 Query Match: 52.33% Indels: 19
 DB: 18 Gaps: 4
 US-10-004-219b-14 (1-455) x AAT50834 (1-1713)
 Qy 2 GlnLeuThrCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPheMet 21
 Db 79 AAACCTGGTCTGCTACTTCCACCACTGGGCCCAAGTACAGACAGGGGAGGCTCGCTTCCTG 138
 Qy 22 ProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArgGln 41
 Db 139 CCCAAGGACTTGGACCCCGCCAGCCCTTTCACCCACCTCACTACGCTTCGCTGGCATGACC 198
 Qy 42 AsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrGlnAlaPheAsnGly 61
 Db 199 AACCAACAGCTGAGCACCACCTGAGTGAATGACGAGACTCTTACAGGAGTTCAATGGC 258
 Qy 62 LeuLysAsnLysAsnSerGlnLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPheGly 81
 Db 259 CTGAAGAAGATGAATCCCAAGCTGAAGACCTGTGTAGCCATCGGAGGCTGGAATTCGGC 318
 Qy 82 ThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThrSer 101
 Db 319 ACTCAGAAGTTTACAGATATGTAGCCAGCGGCCAACACCCGACACCTTTGTCAACTCG 378
 Qy 102 ValIleLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyrPro 121
 Db 379 GCCATCAGGTTTTCGCGCAATACAGCTTTTACGGCCCTTGACTTGGGAGTACCCA 438
 Qy 122 GlySerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValGlnGluMet 141
 Db 439 GGAAGCCAGGGGAGGCCCTCGCGTAGACAAGGAGCGCTTCACACCCCTGTGTAGAGACTTG 498
 Qy 142 ArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThrAla 161
 Db 499 GCCAATGCTTCCAGCAGGAAGCCAGCCACTCAGGGAAGAACGCCCTTCTCTGAGTGCA 558
 Qy 162 AlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSerGln 181

[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 29, 2003, 21:12:38 ; Search time 64.9519 Seconds
(without alignments)
2148.326 Million cell updates/sec

Title: US-10-004-219B-14

Perfect score: 2506
Sequence: 1 YQLTCYFTNQAQYRGLGRF.....QONCOAGLVFTSCDCNNA 455

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame-p2n.model -DEV=xlp
-O=/cgn2_1/USPTO/spool/US10004219/runat_24062003_160229_8184/app_query.fasta_1.1877
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10004219 @CGN 1 1 108 @runat_24062003_160229_8184 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT_DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTCUS COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1323	52.8	1636	US-09-039-198A-1	Sequence 1, Appli
2	1323	52.8	1636	US-08-877-599-1	Sequence 1, Appli
3	1323	52.8	1636	US-09-267-574-1	Sequence 1, Appli
4	1323	52.8	1643	US-08-486-839-3	Sequence 3, Appli
5	1323	52.8	1643	US-09-151-011-3	Sequence 3, Appli
6	1323	52.8	1643	US-09-343-623-3	Sequence 3, Appli
7	1317	52.6	1656	US-09-039-198A-3	Sequence 3, Appli
8	1317	52.6	1656	US-08-877-599-3	Sequence 3, Appli
9	1317	52.6	1656	US-09-267-574-3	Sequence 3, Appli
10	1311.5	52.3	1713	US-08-486-839-5	Sequence 5, Appli
11	1311.5	52.3	1713	US-09-151-011-5	Sequence 5, Appli
12	1311.5	52.3	1713	US-09-343-623-5	Sequence 5, Appli

13	970.5	38.7	1433	1	US-08-694-915-1	Sequence 1, Appli
14	970.5	38.7	1526	1	US-08-694-915-3	Sequence 3, Appli
15	962.5	38.4	1681	5	PCT-US94-07754-4	Sequence 4, Appli
16	957.5	38.2	1681	2	US-08-581-527-4	Sequence 4, Appli
17	863	34.4	2452	2	US-08-524-051-1	Sequence 1, Appli
18	798.5	31.9	1677	4	US-09-545-814-13	Sequence 13, Appli
19	798.5	31.9	1677	4	US-09-545-814-15	Sequence 15, Appli
20	798.5	31.9	1749	4	US-09-545-814-4	Sequence 4, Appli
21	798.5	31.9	1749	4	US-09-545-814-6	Sequence 6, Appli
22	798.5	31.9	2610	4	US-09-545-814-1	Sequence 1, Appli
23	798.5	31.9	2610	4	US-09-545-814-3	Sequence 3, Appli
24	794.5	31.7	1919	4	US-09-545-814-31	Sequence 31, Appli
25	794.5	31.7	1919	4	US-09-545-814-33	Sequence 33, Appli
26	711.5	28.4	1470	4	US-09-292-225-40	Sequence 40, Appli
27	711.5	28.4	1470	4	US-09-292-225-42	Sequence 42, Appli
28	711.5	28.4	1527	4	US-09-292-225-37	Sequence 37, Appli
29	711.5	28.4	1527	4	US-09-292-225-39	Sequence 39, Appli
30	711.5	28.4	1621	4	US-09-292-225-34	Sequence 34, Appli
31	711.5	28.4	1621	4	US-09-292-225-36	Sequence 36, Appli
32	708.5	28.3	1608	4	US-09-292-225-20	Sequence 20, Appli
33	708.5	28.3	1608	4	US-09-292-225-22	Sequence 22, Appli
34	708.5	28.3	1665	4	US-09-292-225-17	Sequence 17, Appli
35	708.5	28.3	1665	4	US-09-292-225-19	Sequence 19, Appli
36	708.5	28.3	1752	4	US-09-292-225-16	Sequence 16, Appli
37	708.5	28.3	1752	4	US-09-292-225-18	Sequence 18, Appli
38	672	26.8	1478	4	US-09-545-814-28	Sequence 28, Appli
39	672	26.8	1478	4	US-09-545-814-30	Sequence 30, Appli
40	472.5	18.9	1167	1	US-07-939-501A-6	Sequence 6, Appli
41	472.5	18.9	1167	4	US-08-448-398-10	Sequence 10, Appli
42	472.5	18.9	1320	1	US-07-939-501A-14	Sequence 14, Appli
43	472.5	18.9	1364	1	US-07-939-501A-13	Sequence 13, Appli
44	472.5	18.9	1405	1	US-07-939-501A-9	Sequence 9, Appli
45	463.5	18.5	1554	1	US-08-045-269C-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-09-039-198A-1
; Sequence 1, Application US/09039198A
; Patent No. 6200951
; GENERAL INFORMATION:
; APPLICANT: Gray, Patrick W.
; TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive/6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/039,198A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 27866/34391
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1636 base pairs
; TYPE: nucleic acid

```

, , STRANDEDNESS: single
, , TOPOLOGY: linear
, , MOLECULE TYPE: CDNA
, , FEATURE:
, , NAME/KEY: CDS
, , LOCATION: 2..1399
, , FEATURE:
, , NAME/KEY: mat_peptide
, , LOCATION: 65..1399
US-09-039-198A-1

```

Alignment Scores:		
Pred. NO.:	6.16e-123	Length:
Score:	132.00	Matches:
Percent Similarity:	76.39%	Conservative:
Best Local Similarity:	52.17%	Mismatches:
Query Match:	52.79%	Indels:
DB:	4	Gaps:
		3
		240
		70
		126
		24
		4
		2
		1636

US-10-004-219B-14 (1-455) x US-09-039-198A-1 (1-1636)

Qy	2	GlnLeuThrCysGlyTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPheMet	21
Db	68	AAACGTGGTCTGTACTTCCACCACTGGGCCCCAGTACAGACAGGGAGGCTCGCTTCCTG	127
Qy	22	ProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArgGln	41
Db	128	CCCAAGGACTTGGAGCCCGAGCTTTGCAACCCACCTCATCTACGCTTCGCTGGCATGACC	187
Qy	42	AsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrGlnAlaPheAsnGly	61
Db	188	AACCACCAGCTGAGCACCACTGAGTGGAAATGACGAGACTCTCTACCAAGGAGTTCAATGCG	247
Qy	62	LeuLysAsnLysAsnSerGlnLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPheGly	81
Db	248	CTGAAGAAGATGAATCCCAAGCTGAAGACCCCTGTTAGCCATCGAGGCTGGAATTTCCGC	307
Qy	82	ThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThrSer	101
Db	308	ACTCAGAGTTTCAAGATATGTAGCCACGCGCCAAACCGTCAGACCTTTGTCAACTCG	367
Qy	102	ValIleLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyrPro	121
Db	368	GCCATCAGGTTTCTGCGCAATACAGCTTTTGACGCGCTTGACCTTGGAGTAGTACCA	427
Qy	122	GlySerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValGlnGluMet	141
Db	428	GGAAGCCAGGGGAGCCCTGCGCGTACAGAAGAGAGCGCTTCAACACCTGGTACAGGACTT	487
Qy	142	ArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThrAla	161
Db	488	GCCATGCTCTCCACAGGAGGCCACAGCTTCAGGGAAGGAAACGCTTCTTCTGAGTGCA	547
Qy	162	AlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSerGln	181
Db	548	GCGGTTCCAGCTGGGCAGACCTATGTGGATCTGGATACGAGGTGGACAAATCGCCAG	607
Qy	182	TyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThr	201
Db	608	AACTCGGATTTTGTCAACCTTATGGCTTACGACTTCCATGGCTCTTGGGAGAGGTCAAG	667
Qy	202	GlyGluAsnSerProLeuTyrLysTyrProThrAspThrGlySerAsnAlaTyrLeuAsn	221
Db	668	GGACATACAGCCCCCTCTACAGAGGCGAAGAGAGTGTGGCAGGACCGACCTCAAC	727
Qy	222	ValAspTyrValIleMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeuIleVal	241
Db	728	GTGGATGCTGCTGCAACAGTGGCTGCAGAAAGGGACCCCTGCCAGCAGCTGATCCTT	787
Qy	242	GlyPheProThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIleGly	261
Db	788	GGCATGCTTACTCTACGAGCGTCTCTTCACTATGGCTCTCTCATCAGACACAGAGTGGG	847

TELEFAX: 312/474-0448

TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1636 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 2..1399

FEATURE:

NAME/KEY: mat_peptide

LOCATION: 65..1399

US-08-877-599-1

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

Length: 1636

Matches: 240

Conservative: 70

Mismatch: 126

Indels: 24

Gaps: 3

US-10-004-219B-14 (1-455) x US-08-877-599-1 (1-1636)

QY 2 GlnLeuThrCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPheMet 21
 DB 68 AAACCTGGTCTGCTACTTACCAACCTGGGCCCCAGTACAGACAGCGGAGGCTCGCTTCCTG 127
 QY 22 ProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArgGln 41
 DB 128 CCCAAGGACTTGGACCCAGCCCTTGGACCCACCTCATCTAGCCCTTCTGCTGGCATGACC 187
 QY 42 AsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrGlnAlaPheAsnGly 61
 DB 188 AACCAACAGCTGAGCACCACCTGAGTGGATGACGAGACTCTTACCAAGAGTTCAATGGC 247
 QY 62 LeuLeuAsnLeuAsnSerGlnLeuIleThrLeuLeuAlaIleGlyTyrAsnPheGly 81
 DB 248 CTGAAGAAGATGAATCCCAAGCTGAAGACCCCTGTAGCCATCGAGGCTGGAAATTCGGC 307
 QY 82 ThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThrSer 101
 DB 308 ACTCAGAACTTACAGATATGTAGCCAGCGGCCACACCGCTCAGACCTTTGTCAACTCG 367
 QY 102 ValIleLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyrPro 121
 DB 368 GGCATCAGGTTCTGGCCAAATACAGCTTGGAGGCTTGGACCTTGGAGTGGAGTACCCA 427
 QY 122 GlySerArgGlySerProGlnAspLysHisLeuPheThrValLeuValGlnGluMet 141
 DB 428 GGAAGCCAGGGAGCCCTCCGCTAGACAGAGGAGCGCTTACCAACCTGTGTACAGGACTTG 487
 QY 142 ArgGluAlaPheGluGlnGluAlaTyrGlnIleAsnLysProArgLeuMetValThrAla 161
 DB 488 GCCAATGCCCTTCAGCAGAGAGCCAGACTCAGGAGAGGAGGCTTCTTCTGAGTGA 547
 QY 162 AlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSerGln 181
 DB 548 GCGGTTCCAGCTGGGCGAGACCTATGTGGATGTGGATACGAGTGGAGCAAAATCGCCAG 607
 QY 182 TyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThr 201
 DB 608 AACCTGGATTTGTCAACCTTATGGCTTACGACTTCCATGGCTCTTGGGAGAGGTCACG 667
 QY 202 GlyGluAsnSerProLeuTyrIleTyrProThrAspThrGlySerAsnAlaTyrLeuAsn 221
 DB 668 GGACATACAGCCCTCTACAGAGGCGAAGAGAGTGGTGCAGCAGCCAGCCTCAAC 727
 QY 222 ValAspTyrValMetAsnTyrTrpTyrAspAsnGlyAlaProAlaGluLysLeuIleVal 241

DB 728 GTGGATGCTGCTGTGCAACAGTGGCTGCAGAGGAGGCCCTTCGCCAGCAAGCTGATCCTT 787
 QY 242 GlyPheProThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIleGly 261
 DB 788 GGCATGCTTACTACGGAGCGCTCTTCACTTGGCTCTCTATCAGACACCAAGAGTGGG 847
 QY 262 AlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyIleTrpAla 281
 DB 848 GCCCAGCCACAGGGCTTGGCAGCTCCAGGCCCTTCCACCAAGGAAGGAGGATGTGGCC 907
 QY 282 TyrTyrGluIleCysThrPheLeuLysAsnGlyAlaThrGlnGlyTrpAspAlaProGln 301
 DB 908 TACTATGAAGTCTGCTCTGG-----AAGGGGGGCCCAACACAGAGAATCCAGATCAG 961
 QY 302 GluValProTyrAlaTyrGlnGlyAsnValTrpValGlyTyrAspAsnIleLysSerPhe 321
 DB 962 AAGTGGCTTACATCTTCCGGGACCAACAGTGGTGGGTGGCTTGTATGATGTGGAGAGCTTC 1021
 QY 322 AspIleLysAlaGlnTrpLeuLysHisAsnLysPheGlyGlyAlaMetValTrpAlaIle 341
 DB 1022 AAAACCAAGGTCAGCTATCTGAAGCAGAAAGGAGCTGGGGGGGCCCATGCTTGGGCACTG 1081
 QY 342 AspLeuAspAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSerThr 361
 DB 1082 GACTTAGATGACTTTTGGCCGCTTCTCTGCAACAGGCGCGATACCCCTCATCCAGACG 1141
 QY 362 LeuLysLysAlaLeu-----GlyLeuGlnSerAlaSerCysThr 374
 DB 1142 CTACGGCAGGAACTGAGTCTTCCATCTTGCCTTCCAGGACCCAGAGCTTGAAGTTCCA 1201
 QY 375 AlaProAlaGlnProIleGluProIleThrAlaAlaProSerGlySerGlyAsnGlySer 394
 DB 1202 AAACCAAGTTCAGCCCTCTGACCT-----1225
 QY 395 GlySerSerSerSerGlySerSerGlySerGlySerGlyPheCysAlaValArgAlaAsn 414
 DB 1226 -----GAGCATGGGCCCGCCAGCCCTGGACAAGACACGTTCTTCCAGGGCAAGCTGAT 1276
 QY 415 GlyLeuTyrProValAlaAsnAsnArgAsnAlaPheTrpHisCysValAsnGlyValThr 434
 DB 1277 GGGCTTATCCCAATCTCCGGGAACGGTCCAGCTTCTACAGCTGTGCAGCGGGGGGGTGTG 1336
 QY 435 TyrGlnGlnAsnCysGlnAlaGlyLeuValPheAspThrSerCysAspCysCysAsnTrp 454
 DB 1337 TTCCAGCAAGAGTGGCCCGCAGAGCGCTGTGTTCACCAACTCTCTGCAAAATGCTGCACCTGG 1396

RESULT 3

US-09-267-574-1

; Sequence 1, Application US/09267574

; Patent No. 6399571

; GENERAL INFORMATION:

; APPLICANT: Gray, Patrick W.

; APPLICANT: Tjoelker, Larry W.

; TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS

; FILE REFERENCE: 27866/35407

; CURRENT APPLICATION NUMBER: US/09/267,574

; CURRENT FILING DATE: 1999-03-12

; EARLIER APPLICATION NUMBER: 09/039,198

; EARLIER FILING DATE: 1998-03-12

; NUMBER OF SEQ ID NOS: 39

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 1

; LENGTH: 1636

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (2) .. (1399)

; FEATURE:

; NAME/KEY: mat_peptide

; LOCATION: (65) .. (1399)

; US-09-267-574-1

Alignment Scores:

Pred. No.: 6,16e-123 Length: 1636
Score: 1323.00 Matches: 240
Percent Similarity: 67.39% Conservative: 70
Best Local Similarity: 52.17% Mismatches: 126
Query Match: 52.79% Indels: 24
DB: 4 Gaps: 3

US-10-004-219B-14 (1-455) x US-09-267-574-1 (1-1636)

```
QY 2 GlnLeuThrCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPheMet 21
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 68 AAATGGTCTGCTACTTCAACAACTGGGCCCATACAGACAGGGGAGGCTGCTTCCTG 127
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::

QY 22 ProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArgGln 41
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 128 CCAAGGACTTGGACCCCGACCTTTGCACCCACCTCATCTACGCCCTTGCCTGCATGACC 187
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::

QY 42 AsnAsnGluIleThrThrIleGluTyrAsnAspValThrLeuTyrGlnAlaPheAsnGly 61
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 188 AACCCAGCTGAGCACCACTGAGTGGATGACGAGCTCTTACACGAGTTCAATGGC 247
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::

QY 62 LeuLysAsnLysAsnSerGlnLeuLysThrLeuLeuAlaIleGlyGlyTyrAsnPheGly 81
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 248 CTGAAGAAGATGAATCCCAAGCTGAAGACCTCTTAGCCATCGGAGCTGGAATTTCCGC 307
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::

QY 82 ThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThrSer 101
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 308 ACTCAGAAGTTTACAGATATGGTAGCGGCGCCCAACCGTCAGACCTTTGTCAACTCG 367
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::

QY 102 ValIleLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTyrGluTyrPro 121
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 368 GCCATCAGGTTCTGCGCAATACAGCTTTGACGGCTTGACCTTGAGTGGAGTACCCA 427
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::

QY 122 GlySerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValGlnGluMet 141
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 428 GGAAGCCAGGGAGCGCTGCGGAGCAAGAGCGCTTCAACACCTTGTGTACAGGACTTG 487
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::

QY 142 ArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThrAla 161
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 488 GCCAATGCCCTTCAGCAGGAAGCCAGACCTCAGGGAAGGAACGCTCTTCTGAGTGCA 547
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::

QY 162 AlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSerGln 181
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 548 GCGTTCAGCTGGCAGACCTATGCGATGCTGGATAGAGTGACAGGAGCAAAATCGCCAG 607
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::

QY 182 TyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTyrGluGlyTyrThr 201
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 608 AACCTGGATTTTGTCAACCTTATGGGCTACGACTTCATGGCTCTTGGGAGAGGTCAAG 667
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::

QY 202 GlyGluAsnSerProLeuTyrLysTyrProThrAspThrGlySerAsnAlaTyrLeuAsn 221
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 668 GGACATAACAGCCCTCTCAAGAGCGCAAGAGAGGTGTGCAGCAGCCAGCCCTCAAC 727
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::

QY 222 ValAspTyrValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeuIleVal 241
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 728 GTGGATGCTGCTGTGCAACAGTGGCTGCAGAGGGGACCCCTGCCAGCAAGCTGATCTT 787
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::

QY 242 GlyPheProThrTyrGlyHisAsnPhelLeuSerAsnProSerAsnThrGlyIleGly 261
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 788 GGCATGCTTACCTTACGAGAGCTCTTCACTGGCTCTCATCAGACACCAAGAGTGGGG 847
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::

QY 262 AlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyIleTyrAla 281
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 848 GCCCCAGCACAGGGTCTGCGCATCTCCAGGCCCTTCCACCAAGGAAGAGGATGCTGGCC 907
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::

QY 282 TyrTyrGluIleCysThrPheLeuLysAsnGlyValThrGlnGlyTyrAspAlaProGln 301
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 908 TACTATGAGTCTGCTCCTCG-----AAGGGGGCCCAACACAGAGATCCAGATCAG 961
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::

QY 302 GluValProTyrAlaTyrGlnGlyAsnValTyrValGlyTyrAspAsnIleLysSerPhe 321
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 962 AAGGTGCCCTTACATCTTCCGGGCAACACCAAGTGGGTGGCTTTTGTATGATGTGGAGCTTC 1021
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
```

```
QY 322 AspIleLysAlaGlnTrpLeuLysHisAsnLysPheGlyGlyAlaMetValTrpAlaIle 341
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1022 AAAACCAAGCTCAGTATCTGAAGCAGAGGACTGGGGGGGCCCATGTCTTGGGCACTG 1081
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::

QY 342 AspLeuAspAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSerThr 361
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1082 GACTTAGTAGTACTTTGGCGGCTTCTCTGCAACACGAGGCGCATACCCCTCATCCAGAG 1141
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::

QY 362 LeuLysLysAlaLeu-----GlyLeuGlnSerAlaSerCysThr 374
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1142 CTACGCGCAGAACTGAGTCTTCATCTGCTTCAGGCACCCAGAGCTTGAAGTTCCA 1201
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::

QY 375 AlaProAlaGlnProIleGluProIleThrAlaAlaProSerGlySerGlyAsnGlySer 394
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1202 AACACAGCTACGCCCTCTGAACCT----- 1225
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::

QY 395 GlySerSerSerSerGlyGlySerSerGlyGlySerGlyPheCysAlaValAlaGln 414
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1226 -----GAGCATGGCCCCAGCCCTGGACAAGACACGTTCTGCCAGGGCAAGCTGAT 1276
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::

QY 415 GlyLeuTyrProValAlaAsnAsnArgAsnAlaPheTrpHisCysValAsnGlyValThr 434
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1277 GGGCTCTATCCCAATCCTCGGAACGGTCCAGCTTCTACAGCTGTGCAGCGGGCGGCTG 1336
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::

QY 435 TyrGlnGlnAsnCysGlnAlaGlyLeuValPheAspThrSerCysAspCysCysAsnTrp 454
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1337 TTCCAGCAAGGTGCCGACAGCGCTGTGTGTTTCAGCAACTCTGCAATGCTGCACCTGG 1396
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::

RESULT 4
US-08-486-839-3
; Sequence 3, Application US/08486839
; Patent No. 5928928
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: A human chitinase, its recombinant
; TITLE OF INVENTION: production, its use for decomposing chitin, its use
; TITLE OF INVENTION: in therapy or prophylaxis against infection diseases.
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11758
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,839
; FILING DATE: 07 - June - 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Baron, Ronald J.
; REGISTRATION NUMBER: 29,281
; REFERENCE/DOCKET NUMBER: 294-26
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1643 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; US-08-486-839-3
```

Alignment Scores:

Pred. No.: 6.2e-123 Length: 1643
 Score: 1323.00 Matches: 240
 Percent Similarity: 67.39% Conservative: 70
 Best Local Similarity: 52.17% Mismatches: 126
 Query Match: 52.79% Indels: 24
 DB: 2 Gaps: 3

US-10-004-219B-14 (1-455) x US-08-486-839-3 (1-1643)

QY 2 GlnLeuThrCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPheMet 21
 DB 79 AAACCTGGTCTGCTACTTACCAACTGGGCCAGTACAGACAGGGAGGCTCGCTTCCTG 138
 QY 22 ProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArgGln 41
 DB 139 CCAAGGACTTGGACCCAGCTTTGACCCACCTCATCTACGCTTGTGCTGGCATGCC 198
 QY 42 AsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrGlnAlaPheAsnGly 61
 DB 199 AACCAACGACTGAGCACCACTGAGTGGATGACGAGACTCTTACCGAGGATTCATGGC 258
 QY 62 LeuLysAsnLysAsnSerGlnLeuLysThrLeuAlaIleGlyGlyTrpAsnPheGly 81
 DB 259 CTGAAGAAGATGATCCCAAGCTGAAGCCCTGTAGCCATCGAGGCTGGGATTTGGGC 318
 QY 82 ThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThrSer 101
 DB 319 ACTCAGAAGTTCACAGATATGTTAGCCAGCGGCCCAACACCGTCAGACCTTTGTCAACTCG 378
 QY 102 ValIleLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyrPro 121
 DB 379 GCCATCAGGTTCTGCGCAATACAGCTTGTGAGCTTGTGAGCTTGTGAGTACCCA 438
 QY 122 GlySerArgGlySerProGlnAspLysHisLeuPheThrValLeuValGlnGluMet 141
 DB 439 GGAACCCAGGGAGCCCTCCCTAGACAGAGAGCGCTTCACACCTGTGTGACGACTTG 498
 QY 142 ArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThrAla 161
 DB 499 GCCAATGCTTCCAGCAGGAGGCCAGCCTCAGGGAGGAGCGCTTCTTCTGAGTGCA 558
 QY 162 AlaValAlaAlaGlyLysSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSerGln 181
 DB 559 GCGGTTCCAGCTGGGAGCAGCATGTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 618
 QY 182 TyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThr 201
 DB 619 AACCTGGATTTGTCAACCTTATGGCTACGACTTCCATGGCTCTTGGGAGAGGTGACG 678
 QY 202 GlyGluAsnSerProLeuTyrLysTyrProThrAspThrGlySerAsnAlaTyrLeuAsn 221
 DB 679 GGACATAACAGCCCTCTACAGAGGCAAGAGAGAGTGGTGACGACGACGCTCAAC 738
 QY 222 ValAspTyrValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeuVal 241
 DB 739 GTGGATGCTGCTGTGCAACAGTGGCTGCAAGAGGGGAGCCCTGCCAGCAAGCTGATCCTT 798
 QY 242 GlyPheProThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIleGly 261
 DB 799 GGCATGCTTACCTAGGAGCGCTTCAACCTGCGCTCTCATCATCAGACACGAGGTGGGG 858
 QY 262 AlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyIleTrpAla 281
 DB 859 GCGCCAGCAGAGGCTGCGCACTCCAGCCCTTCCACCAAGAGAGGGAGGTGCGCC 918
 QY 282 TyrTyrGluIleCysThrPheLeuLysAsnGlyAlaThrGlnGlyTyrAspAlaProGln 301
 DB 919 TACTATGAAGTCTGCTCCTGG-----AAGGGGGGCCAACCAACAGAGAATCCAGGATCAG 972
 QY 302 GluValProTyrAlaTyrGlnGlyAsnValTrpValGlyTyrAspAsnIleLysSerPhe 321
 DB 973 AAGGTGCCCTATCTTCCGGGACCAACCGAGTGGGTGGGCTTTGATGATGTGGAGAGCTTC 1032

QY 322 AspIleLysAlaGlnTrpLeuLysHisAsnLysPheGlyGlyAlaMetValTrpAlaIle 341
 DB 1033 AAAACCAAGGTGAGCTATCTGAAGCAGAGGAGTGGCGGGCCATGCTCTGGGCACTG 1092
 QY 342 AspLeuAspAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSerThr 361
 DB 1093 GACTTAGATGACTTTGGCGGCTTCTCTGCAACAGGCGCCGATACCCCTCATCCAGACG 1152
 QY 362 LeuLysLysAlaLeu-----GlyLeuGlnSerAlaSerCysThr 374
 DB 1153 CTACGGCAGGAGTACTGCTTCCATCTTCCATCTTCCAGCACCCAGAGCTTGAAGTTCCA 1212
 QY 375 AlaProAlaGlnProIleGluProIleThrAlaAlaProSerGlySerGlyAsnGlySer 394
 DB 1213 AAACGAGTCCAGCCCTCTGAACCT----- 1236
 QY 395 GlySerSerSerSerGlyGlySerSerSerGlyGlySerGlyPheCysAlaValArgAlaAsn 414
 DB 1237 -----GAGCATGGCGCCCGCCAGCCCTCGCAAGACAGCTTCTGCCAGGGCAAGCTGAT 1287
 QY 415 GlyLeuTyrProValAlaAlaAsnAsnArgAsnAlaPheTrpHisCysValAsnGlyValThr 434
 DB 1288 GGGCTCTATCCCAATCTCTCGGGAGCGTCCAGCTTCTACAGCTGTGCAGCGGGCGGCTG 1347
 QY 435 TyrGlnGlnAsnCysGlnAlaGlyLeuValPheAspThrSerCysAspCysAsnTrp 454
 DB 1348 TTCAGCAAGCTGCGCCGACAGGCGCTGTGTTTCAGCAACTCTCTGCAATGCTGCACCTGG 1407

RESULT 5

US-09-151-011-3
 ; Sequence 3, Application US/09151011
 ; Patent No. 6057142

GENERAL INFORMATION:

APPLICANT:
 TITLE OF INVENTION: A Human Chitinase, Its Recombinant
 TITLE OF INVENTION: Production, Its Use For Decomposing Chitin, Its Use in
 NUMBER OF SEQUENCES: 16
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Hoffmann & Baron, LLP
 STREET: 6900 Jericho Turnpike
 CITY: Syosset
 STATE: New York
 COUNTRY: United States of America
 ZIP: 11791

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/151,011
 FILING DATE: 10 - September - 1998
 ATTORNEY/AGENT INFORMATION:
 NAME: Morris, Robert C.
 REGISTRATION NUMBER: 42,910
 REFERENCE/DOCKET NUMBER: 294-32 DIV
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (516) 822-3550
 TELEFAX: (516) 822-3582
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1643 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: unknown
 TOPOLOGY: unknown
 MOLECULE TYPE: cDNA
 HYPOTHETICAL: NO
 US-09-151-011-3

Alignment Scores: 6.2e-123 Length: 1643
 Pred. No.:

Score: 1323.00 Matches: 240
Percent Similarity: 67.39% Conservative: 70
Best Local Similarity: 52.17% Mismatches: 126
Query Match: 52.79% Indels: 24
DB: 3 Gaps: 3

US-10-004-219B-14 (1-455) x US-09-151-011-3 (1-1643)

QY 2 GlnLeuThrCysThrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPheMet 21
DB 79 AAACGTGGTCTGCTACTTACCAACTGGGCCAGTACAGACAGGGGAGGCTCGCTTCCTG 138
QY 22 ProAspAsnLeuAspProCysLeuCysThrHisLeuLeuTyrAlaPheAlaGlyArgGln 41
DB 139 CCCAGGAGCTGGACCCAGCCTTTGCACCCACCTCATCTAGCCCTTCGCTGGCATGACC 198
QY 42 AsnAsnGluLeuThrThrileGluTrpAsnAspValThrLeuTyrGlnAlaPheAsnGly 61
DB 199 AACCCACGCTGAGCACCAGTGTAGTGGATGAGACGACTCTTACCAGGAGTTCAATGGC 258
QY 62 LeuLysAsnLysAsnSerGlnLeuLysThrLeuLeuAlaileGlyGlyTrpAsnPheGly 81
DB 259 CTGAAGAAGATGATCCCAAGCTGAGACCCCTGTAGCCATCGGAGCTGGAAATTCGGC 318
QY 82 ThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheileThrSer 101
DB 319 ACTCAGAAGTTTACAGATATGTTAGTGGTGGCCCAACACCGTTCAGACCTTTGTCAACTCG 378
QY 102 ValileLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyrPro 121
DB 379 GCATCAGGTTTCTGGCAAAATACAGCTTTGACGGCTTGTACCTTGGAGTACCCA 438
QY 122 GlySerArgGlySerProGlnAspLysHisLeuPheThrValLeuValGlnGluMet 141
DB 439 GGNAGCCAGGGAGCCCTCCGTAGACAGAGCGCTTCACACCTCTGTACAGACTTG 498
QY 142 ArgGluAlaPheGluGlnGluAlaLysGlnileAsnLysProArgLeuMetValThrAla 161
DB 499 GCCAATGCTCTCCAGCAGGAAGCCAGACCTCAGGGAAGAAAGCGCTTCTTCTAGTGA 558
QY 162 AlaValAlaAlaGlyLeuSerAsnileGlnSerGlyTyrGluileProGlnLeuSerGln 181
DB 559 GCGTTCAGCTGGCAGACCTATGATGGATGCTGGATGAGTGACAAATTCGCCAG 618
QY 182 TyrLeuAspTyrileHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThr 201
DB 619 AACCTGGATTTGTCACCTTATGGCTACGACTTCCATGGCTCTTGGGAGAGGTCA 678
QY 202 GlyGluAsnSerProLeuTyrLysTyrProThrAspThrGlySerAsnAlaTyrLeuAsn 221
DB 679 GGACATAACAGCCCTCTACAGAGGCAAGAGAGAGTGGTGCAGCAGCCAGCTCAAC 738
QY 222 ValAspTyrValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeuileVal 241
DB 739 GTGGATGCTGTGTCAACAGTGGCTGCAAGAGGGGACCCCTGCCAGAGCTGATCCTT 798
QY 242 GlyPheProThrTyrGlyHisAsnPhelleLeuSerAsnProSerAsnThrGlyileGly 261
DB 799 GGCATGCTTACCTACGAGCGCTCTTACACTGGCTCTCATCACACAGAGTGGGG 858
QY 262 AlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyileTrpAla 281
DB 859 GCCCCAGCCACAGGCTGTGGCACTCCAGGCCCTTCCACCAAGGAGGAGGATGCTGGCC 918
QY 282 TyrTyrGluileCysThrPheLeuLysAsnGlyAlaThrGlnGlyTrpAspAlaProGln 301
DB 919 TACTATGAGTGTGCTCTGTGG-----AGGGGGGCCCAACACAGAGATCCAGGATCAG 972
QY 302 GluValProTyrAlaTyrGlnGlnValTrpValGlyTyrAspAsnileLysSerPhe 321
DB 973 AAGTGCCTTACATCTCCGGGACACCACTGAGTGGTGGCTTTGATGATGTGGAGCTTC 1032
QY 322 AspIleLysAlaGlnTrpLeuLysHisAsnLysPheGlyGlyAlaMetValTrpAlaile 341

DB 1033 AAAACCAAGGTGAGTATCTGAAGCAGAGGAGTGGGGGGGCCCATGCTCTGGGCACTG 1092
QY 342 AspLeuAspAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuileSerThr 361
DB 1093 GACTTAGATGACTTGGCCGGCTTCTCCTGCAACAGGCGCCGATACCCCTCATCCAGAG 1152
QY 362 LeuLysLysAlaLeu-----GlyLeuGlnSerAlaSerCysThr 374
DB 1153 CTACGGCAGGAACCTGAGTCTTCATACTTGCCTTCAGGCACCCAGAGAGTGAAGTTCCA 1212
QY 375 AlaProAlaGlnProileGluProileThrAlaAlaProSerGlySerGlyAsnGlySer 394
DB 1213 AAACAGGTGACCCCTCTGAACCT----- 1236
QY 395 GlySerSerSerGlyGlySerSerGlyGlySerGlyPheCysAlaValAlaArgAlaAsn 414
DB 1237 -----GAGCATGGCCCCCAGCCCTGCAGACAGACGTTCTGCAGGGCAAGCTGAT 1287
QY 415 GlyLeuTyrProValAlaAsnAsnArgAsnAlaPheTrpHisCysValAsnGlyValThr 434
DB 1288 GGGCTCTATCCCAATCCTCGGGAAGCGTCCAGCTTCTACAGCTGTGCAGCGGGGGCTG 1347
QY 435 TyrGlnGlnAsnCysGlnAlaGlyLeuValPheAspThrSerCysAspCysCysAsnTrp 454
DB 1348 TTCCAGCAAGAGCTGCCGACAGGCGCTGGTGTTCAGCAACTCTGCAAAATGCTGCACCTGG 1407
RESULT 6
US-09-343-623-3
Sequence 3, Application US/09343623
Patent No. 6303118
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: A human chitinase, its recombinant
TITLE OF INVENTION: production, its use for decomposing chitin, its use
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSER: Hoffmann & Baron
STREET: 350 Jericho Turnpike
CITY: Jericho
STATE: New York
COUNTRY: United States of America
ZIP: 11758
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/343,623
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/486,839
FILING DATE: 07-June-1995
ATTORNEY/AGENT INFORMATION:
NAME: Baron, Ronald J.
REGISTRATION NUMBER: 29,281
REFERENCE/DOCKET NUMBER: 294-26
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 822-3550
TELEFAX: (516) 822-3582
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1643 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
US-09-343-623-3

Alignment Scores:

Pred. No.: 6,2e-123 Length: 1643
 Score: 1323.00 Matches: 240
 Percent Similarity: 67.3% Conservative: 70
 Best Local Similarity: 52.17% Mismatches: 126
 Query Match: 52.7% Indels: 24
 DB: 4 Gaps: 3

US-10-004-219B-14 (1-455) x US-09-343-623-3 (1-1643)

QY	2	GlnLeuThrCysThrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPheMet	21
DB	79	AAACTGGTCTGCTACTTACCAACCTGGGCCAGTACAGACAGGGGAGGCTCGCTTCCTG	138
QY	22	ProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArgGln	41
DB	139	CCCAAGGACTTGGACCCCGAGCTTTGCACCCACCTCATCTAGCCCTTCCTGGCATGCC	198
QY	42	AsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrGlnAlaPheAsnGly	61
DB	199	AACCAACCACTGAGCACCACCTGAGTGAATGACGAGACTCTTACCAAGAGTTCAATGCG	258
QY	62	LeuTyrAsnLysAsnSerGlnLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPheGly	81
DB	259	CTGAAGAAGATGAATCCCAAGCTGAAGACCTGTGTAGCCATCGGAGGCTGGAAATTCGCG	318
QY	82	ThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThrSer	101
DB	319	ACTCAGAATTCACAGATATGTTAGCCAGGCCCAACACCGTCAGACCTTTGTCACTCG	378
QY	102	VallLeuLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyrPro	121
DB	379	GCCATCAGCTTCTCGCAAAATACAGCTTTCAGCGCTTGACCTTGAGCTGGGAGTACCA	438
QY	122	GlySerArgGlySerProGlnAspLysHisLeuPheThrValLeuValGlnGluMet	141
DB	439	GGAAGCCAGGGAGGCGCTCGCTGACAGGAGCGCTTCACACCTCGGTACAGAGCTG	498
QY	142	ArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThrAla	161
DB	499	GCCATGCTTCCAGCAGGAGCCAGACCTCAGGAGGAGAACGCTTCTTCTGAGTGCA	558
QY	162	AlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSerGln	181
DB	559	GCGGTTCAGCTGGCAGACCTATGTGGATGCTGTGATGAGGTGACGAGGTGACAAAATGCC	618
QY	182	TyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThr	201
DB	619	AACCTGGATTTTGTCAACCTTATGGCTACGACTTCCATGGCTCTTGGGAGAGTCCAG	678
QY	202	GlyGluAsnSerProLeuTyrLysTyrProThrAspThrGlySerAsnAlaTyrLeuAsn	221
DB	679	GGACATACAGCCCTCTACAGAGGCGAGAGAGAGTGTGTGACGAGCCAGCCTCAAC	738
QY	222	ValAspTyrValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeuIleVal	241
DB	739	GTGGATGCTGCTGTCAACAGTGGCTGCAAGAGGAGCCCTGCCAGCAAGCTGATCTT	798
QY	242	GlyPheProThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIleGly	261
DB	799	GGCATGGCTTACCTACGAGCGCTCTTCACCTGAGCTTCCATGGCTCTTCATCAGAC	858
QY	262	AlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyIleTrpAla	281
DB	859	GCCCAAGCAGCAGGCTGTGGCACTCAGGCCCTTTCACCAAGAGAGGAGGAGTGTGGCC	918
QY	282	TyrTyrGluIleCysThrPheLeuLysAsnGlyAlaThrGlnGlyTrpAspAlaProGln	301
DB	919	TACTATGAAGTCTGCTCCTGG-----AAGGGGGCCCAACACAGAGAATCCAGGATCAG	972
QY	302	GluValProTyrAlaTyrGlnGlyAsnValTrpValGlyTyrAspAsnIleLysSerPhe	321
DB	973	AAGGTGCCCTATCTTCCGGACAAACAGTGGGTGGGCTTTGATGATGTGGAGAGCTTC	1032

QY	322	AspIleLysAlaGlnTrpLeuLysHisAsnLysPheGlyGlyAlaMetValTrpAlaIle	341
DB	1033	AAACCAAGGTGAGCTATCTGAAGCAGAGGAGGACTGGGCGGGCCCATGCTCTGGGCACTG	1092
QY	342	AspLeuAspAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSerThr	361
DB	1093	GACTTAGATGACTTTTGGCGGCTTCTCTGCAACCAAGGCGCGATACCCCTCATCCAGAGC	1152
QY	362	LeuLysLysAlaLeu-----GlyLeuGlnSerAlaSerCysThr	374
DB	1153	CTAGGCGAGGAAGTGAAGTCTTCATCTTGGCTTTCAGGACCCCGAGAGCTTGAAGTTCCA	1212
QY	375	AlaProAlaGlnProIleGluProIleThrAlaAlaProSerGlySerGlyAsnGlySer	394
DB	1213	AAACCAAGTCAAGCCCTCTGAACCT-----	1236
QY	395	GlySerSerSerSerGlyGlySerSerGlyGlySerGlyPheCysAlaValArgAlaAsn	414
DB	1237	-----GAGCATGGGCCCGCCCTGCAACAGACAGTTCCTGCGAGGCAAGCTGAT	1287
QY	415	GlyLeuTyrProValAlaAsnAsnArgAsnAlaPheTrpHisCysValAsnGlyValThr	434
DB	1288	GGGCTCTATCTCCCAATCTCTCGGGAACGGTCCAGCTTCTACAGCTGTGCGCGGGCGCTG	1347
QY	435	TyrGlnGlnAsnCysGlnAlaGlyLeuValPheAspThrSerCysAspCysCysAsnTrp	454
DB	1348	TTCCAGCAAGAGTCCCGCAGCAGGCTGTGTTCAGCAACTCTTGCACACTGCTGCACCTGG	1407

RESULT 7

US-09-039-198A-3
 ; Sequence 3, Application US/09039198A
 ; Patent No. 6200951
 ; GENERAL INFORMATION:
 ; APPLICANT: Gray, Patrick W.
 ; APPLICANT: Tjoelker, Larry W.
 ; TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
 ; NUMBER OF SEQUENCES: 34
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
 ; STREET: 233 South Wacker Drive/6300 Sears Tower
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: United States of America
 ; ZIP: 60606-6402
 ; COMPUTER READABLE FORM:
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/039,198A
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Rin-Laures, Li-Hsien
 ; REGISTRATION NUMBER: 33,547
 ; REFERENCE/DOCKET NUMBER: 27866/34391
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (312) 474-6300
 ; TELEFAX: (312) 474-0448
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1656 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 27..1424
 ; FEATURE:
 ; NAME/KEY: mat_peptide

LOCATION: 90..1424
US-09-039-198A-3

Alignment Scores:

Pred. No.: 2, 5e-122 Length: 1656
Score: 1317.00 Matches: 239
Percent Similarity: 67.17% Conservative: 70
Best Local Similarity: 51.96% Mismatches: 127
Query Match: 52.55% Indels: 24
DB: 4 Gaps: 3

US-10-004-219B-14 (1-455) x US-09-039-198A-3 (1-1656)

```
QY 2 GlnLeuThrCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPheMet 21
DB 93 AAACGTGGTCTGCTACTTCCACCACTGGGCCAGTACAGACAGGGGAGGCTCGCTTCCTG 152
QY 22 ProAspAsnLeuAspProCysLeuCysThrHisLeuLeuLeuLeuLeuLeuLeuLeuLeu 41
DB 153 CCCAAGGACTTGGACCCAGCCTTTGCACCCACCTCATCTAGCCCTTCGCTGGCATGACC 212
QY 42 AsnAsnGluLeuThrThrileGluTrpAsnAspValThrLeuTyrGlnAlaPheAsnGly 61
DB 213 AACCCAGCTAGCAGCACTAGTGGATGAGAGACTCTTACCAGAGATTCAATGGC 272
QY 62 LeuLysAsnLysAsnSerGlnLeuLysThrLeuLeuAlaileGlyGlyTrpAsnPheGly 81
DB 273 CTGAAGAAGTGAATCCCAAGCTGAGAGCCCTGTAGCCATCGGAGCTTGAATTTTCAGC 332
QY 82 ThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheleThrSer 101
DB 333 ACTCAAGAGTTTCACAGATATGTTAGTGGAGCCGCAACACCGTCAGACCTTTGTCAACTCG 392
QY 102 ValileLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyrPro 121
DB 393 GCATCAGGTTTCTGGCAAAATACACTTTCAGCGCTTACCTTGCATGGAGTACCA 452
QY 122 GlySerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValGlnGluMet 141
DB 453 GGAAGCCAGGGAGCCCTCGCTAGACAAAGGAGCGCTTCACAACTGGTACAGGACTG 512
QY 142 ArgGluAlaPheGluGlnGlnAlaLysGlnileAsnLysProArgLeuMetValThrAla 161
DB 513 GCCAATGCTCTCCAGCAGGAAGCCCACTCAGGGAAGGAAAGCGCTTCTTGTAGTGCA 572
QY 162 AlaValAlaAlaGlyLeuSerAsnileGlnSerGlyTyrGluLeuProGlnLeuSerGln 181
DB 573 GCGGTTCCAGTGGGAGACCTATGTTGGATGCTGGATAGAGTGGACAAATCGCCAG 632
QY 182 TyrLeuAspTyrileHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThr 201
DB 633 AACCTGGATTTTGTCAACTTATGGCTACGACTTCCATGGCTTTGGGAGAGGTTCAG 692
QY 202 GlyGluAsnSerProLeuTyrLysTyrProThrAspThrGlySerAsnAlaTyrLeuAsn 221
DB 693 GGACATAACAGCCCTCTACAAGAGGCAAGAGAGTGGTGCAGCAGCCAGCTCAAC 752
QY 222 ValAspTyrValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeuileVal 241
DB 753 GTGGATGCTGTGCAACAGTGGCTGCGAGAGGGAGCCCTGCCAGCAAGCTGATCCTT 812
QY 242 GlyPheProThrTyrGlyHisAsnPhelileLeuSerAsnProSerAsnThrGlyileGly 261
DB 813 GGCATGCTTACCTACGAGCGCTCTTACACTGGCTCTCATCAGACACAGAGTGGGG 872
QY 262 AlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyileTrpAla 281
DB 873 GCCCCAGCACAGGCTGCGCACTCCAGGCCCTTCAACAAAGGAGGAGGAGTGTGGCC 932
QY 282 TyrTyrGluileCysThrPheLeuLysAsnGlyAlaThrGlnGlyTrpAspAlaProGln 301
DB 933 TACTATGAGTCTGCTCTGG-----AAGGGGGGCCCAACACAGAGAATCCAGGATCAG 986
```

```
QY 302 GluValProTyrAlaTyrGlnGlyAsnValTrpValGlyTyrAspAsnlelySerPhe 321
DB 987 AAGTGGCCCTACATCTCGGACCAACACAGTGGTGGCTTTGATCATGTGGAGAGCTTC 1046
QY 322 AspIleLysAlaGlnTrpLeuLysHisAsnLysPheGlyGlyAlaMetValTrpAlaile 341
DB 1047 AAACCAAGAGTTCAGTATCTGAAGCAGAGGAGTGGGGCGGGCCATGTCTGGGCACTG 1106
QY 342 AspLeuAspAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuileSerThr 361
DB 1107 GACTTAGATGACTTTCGGGCTTCTCTGCAACAGGCGCCGATACCCCTCATCCAGAGC 1166
QY 362 LeuLysLysAlaLeu-----GlyLeuGlnSerAlaSerCysThr 374
DB 1167 CTACGGCAGGAACCTGAGTCTTCATACATCTTCAGGCACCCAGAGCTTGAAGTTCCA 1226
QY 375 AlaProAlaGlnProIleGluProIleThrAlaAlaProSerGlySerGlyAsnGlySer 394
DB 1227 AAACAGGTTCAGCCCTCTGAACCT-----1250
QY 395 GlySerSerSerSerGlyGlySerSerGlyGlySerGlyPheCysAlaValArgAlaAsn 414
DB 1251 -----GACCATGGGCCCGCCAGCTTGCAGGCGCAAGACGTTCCAGGCGCAAGCTGAT 1301
QY 415 GlyLeuTyrProValAlaAsnAsnArgAsnAlaPheTrpHisCysValAsnGlyValThr 434
DB 1302 GGGCTCTATCCCAATCTCTCGGAACGGTCCAGCTTCTACAGCTGTGCAGCGGGGGCTG 1361
QY 435 TyrGlnGlnAsnGlyLeuValPheAspThrSerCysAspCysCysAsnTrp 454
DB 1362 TTCCAGCAAGAGTCCCGCAGCGCTGTGTGTTTCAGCAACTCTCTGCAAAATGCTGCACCTGG 1421
```

RESULT 8

US-08-877-599-3

Sequence 3, Application US/08877599

Patent No. 6372212

GENERAL INFORMATION:

APPLICANT: Gray, Patrick W.

TITLE OF INVENTION: Chitinase Materials and Methods

NUMBER OF SEQUENCES: 17

CORRESPONDENCE ADDRESS:

ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun

STREET: 6300 Sears Tower, 233 South Wacker Drive

CITY: Chicago

STATE: Illinois

COUNTRY: United States of America

ZIP: 60606-6402

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/877,599

FILING DATE:

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/663,618

FILING DATE: 14-JUN-1996

ATTORNEY/AGENT INFORMATION:

NAME: Rin-Laures, Li-Hsien

REGISTRATION NUMBER: 33,547

REFERENCE/DOCKET NUMBER: 27866/33994

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312/474-6300

TELEFAX: 312/474-0448

TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 1656 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

Pred. No.	2.5e-122	Length:	1656
Alignment Scores:	Score:	Matches:	239
	67.17%	Conservative:	70
	Percent Similarity:	Mismatches:	127
	Best Local Similarity:	Indels:	24
	51.96%	Gaps:	3
	Query Match:		
	52.55%		
DB:	4		

QY	362	LeuYsGlyAlaLeuGlyLeuGlnSerAlaSerCysThrAlaProAlaGlnPro-IleG1	381
DB	1153	CTACGCGAGCAACTGAATGGGTAAAGCCCTTAACCTGCTGTGCATCATGTGAGGCCAGGTGTT	1212
QY	381	uProIleThr-AlaAla-----	392
DB	1213	GCTGTGGCACTGTGCTTCAGCTGTAGTCTTCCATACATTGCGCTTCAGGCACCCAGAGC	1272
QY	393	-----GlySerGlySerSerSerSerGlyGlySerSerGlyGlySerG	407
DB	1273	TTGAAGTTCCAAAACACAGGTCAGCCCTCTGAACCTGAGCATGSCCAAGCCCTGGACAAG	1332
QY	407	ly---PheCysAlaValArgAlaAsnGlyLeuTyPrProValAlaAlaAsnAsnArgAsnAlap	426
DB	1333	ACAGGTTCTGCCAGGCAAGCTGTATGGGGCTCTATCCCAATCTCTCGGGAACGGTCCAGCT	1392
QY	426	hetrPhisCysValAsnGlyValThrTyGlnGlnAsnCysGlnAlaGlyLeuValPheA	446
DB	1393	TCTACAGCTGTGCAGCGGGGGCGCTGTTCCAGCAAAGCTGCCCGACAGGCGCTGTTGTTCA	1452
QY	446	spThrSerCysAspCysCysAsnTyr	454
DB	1453	GCAACTCTGCAAAATGCTGCACCTGG	1478

RESULT 11

US-09-151-011-5

Sequence 5, Application US/09151011

Patent No. 6057142

GENERAL INFORMATION:

APPLICANT:

TITLE OF INVENTION: A Human Chitinase, Its Recombinant

TITLE OF INVENTION: Production, Its Use For Decomposing Chitin, Its Use i

TITLE OF INVENTION: Therapy or Prophylaxis Against Infection Diseases.

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: Hoffmann & Barton, LLP

STREET: 6900 Jericho Turnpike

CITY: Syosset

STATE: New York

COUNTRY: United States of America

ZIP: 11791

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

SOFTWARE:

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/151,011

FILING DATE: 10 - September - 1998

ATTORNEY/AGENT INFORMATION:

NAME: Morris, Robert C.

REGISTRATION NUMBER: 42,910

REFERENCE/DOCKET NUMBER: 294-32 DIV

TELECOMMUNICATION INFORMATION:

TELEPHONE: (516) 822-3550

TELEFAX: (516) 822-3582

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 1713 base pairs

TYPE: nucleic acid

STRANDEDNESS: unknown

TOPOLOGY: unknown

MOLECULE TYPE: cdna

HYPOTHETICAL: NO

US-09-151-011-5

Alignment Scores:		
Pred. No.:	9,37e-122	Length: 1713
Score:	1311.50	Matches: 244
Percent Similarity:	68.09%	Conservative: 76
Best Local Similarity:	51.91%	Mismatches: 131

Length:	1713
Matches:	244
Conservative:	76
Mismatches:	131
Indels:	19

DB: 3 Gaps: 4
US-10-004-219b-14 (1-455) x US-09-151-011-5 (1-1713)
QY 2 GlnLeuThrCysTyrPheThrSerThrAlaGlnTyrArgProGlyLeuGlyArgPheMet 21
DB 79 AAACCTGGTCTGCTACTTACCACTACCACTGGGCGCCAGTACAGACAGGGGAGGCTGCTTCTG 138
QY 22 ProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArgGln 41
DB 139 CCCAAGGATGTGACCCAGCCCTTTCACCCACCTCATCTACCCCTTCCTGCGCATGACC 198
QY 42 AsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrGlnAlaPheAsnGly 61
DB 199 AACCCAGCTGAGCACCACCTGAGTGGATGACGAGACTCTCTACCCAGGATTCATATGGC 258
QY 62 LeuLysAsnLysAsnSerGlnLeuLysThrLeuLeuAlaIleGlyClyTrpAsnPheGly 81
DB 259 CTGAAGAAGATGAATCCCAAGCTGAAGACCTGTTAGCCATCGAGGCTGGAATTTCCGC 318
QY 82 ThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThrSer 101
DB 319 ACTCAGAAGTTACAGATATGTAGCCACGCGCCACACCGTCAGACCTTTGTCAACTCG 378
QY 102 ValIleLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyrPro 121
DB 379 GCCATCAGTTCTCGGCAATACAGCTTTGACGGCTTGACCTGACTGGGAGTACCCA 438
QY 122 GlySerArgGlySerProGlnAspLysHisLeuPheThrValLeuValGlnGluMet 141
DB 439 GGAAGCCAGGGAGCCCTCGCGTAGACAGGAGCGCTTCACACCCCTGGTACAGACTTG 498
QY 142 ArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThrAla 161
DB 499 GCCAATGCTCTCAGCAGGAGCCCACTCAGGAGAGAGCGCTTCTTCAGTGCA 558
QY 162 AlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSerGln 181
DB 559 GCGTTCCAGCTGGCAGACCTATGTGGATGCTGGATACGAGTGGACAAATCGCCAG 618
QY 182 TyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThr 201
DB 619 AACCTGGATTTGTCAACCTTATGGCTCAGACTCCATGGCTCTTGGGAGAGGTGACG 678
QY 202 GlyGluAsnSerProLeuTyrIleTyrProThrAspThrGlySerAsnAlaTyrLeuAsn 221
DB 679 GGCATACAGCCCTCTCACAAGAGCGCAGAGAGAGTGTCCAGCAGCCAGCTCAAC 738
QY 222 ValAspTyrValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeuIleVal 241
DB 739 GTGGATGCTGCTGCAACAGTGGCTGCAGAGGGGACCCCTGCCAGCAAGCTGATCCTT 798
QY 242 GlyPheProThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIleGly 261
DB 799 GGCATGCTTACCTACGAGCGCTCTTCACTGGCTCTCTCATCAGACACAGAGTGGGG 858
QY 262 AlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyIleTyrAla 281
DB 859 GCCCCAGCCAGCGGTCTGGCACTCAGGCCCTTCCACAGAGAGAGGAGGTGGCC 918
QY 282 TyrTyrGluIleCysThrPheLeuLysAsnGlyAlaThrGlnGlyTrpAspAlaProGln 301
DB 919 TACTATGAAGTCTGCTCTGG-----AAGGGGGCCACCAACAGAGAATCCAGGATCAG 972
QY 302 GluValProTyrAlaTyrGlnGlyAsnValTrpValGlyTyrAspAsnIleLysSerPhe 321
DB 973 AAGGTGCCCTTACATCTTCCGGGACACCACTGAGTGGGTGGCTTTGATGATGTGGAGCTTC 1032
QY 322 AspIleLysAlaGlnTrpLeuLysHisAsnLysPheGlyClyAlaMetValTrpAlaIle 341
DB 1033 AAACACAGGTGAGTATCTGAAGCAGAGAGGAGTGGGGGGCCCATGTCTGGGCACTG 1092
QY 342 AspLeuAspAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSerThr 361

DB 1093 GACTTAGTAGTACTTTGGCGGCTTCTCTGCAACGAGGCGGATACCCCTCATCCAGACG 1152
QY 362 LeuLysLysAlaLeuGlyLeuGlnSerAlaSerCysThrAlaProAlaGlnPro-IleG1 381
DB 1153 CTACGGCAGGAACCTGAATGGGTAAAGCCTTAACCTGCTCATGTGAGGCCAGGTGT 1212
QY 381 uProIleThr-AlaAla-----ProSerGlySerGlyAsn- 392
DB 1213 GCCTGTGGCAGCTGTCTTCAGCTGTAGTCTTCCATACCTTGCCTTCAGGACCCCGAGAGC 1272
QY 393 -----GlySerGlySerSerSerSerGlyGlySerSerGlyGlySerG 407
DB 1273 TTGAAGTTCCAAACACGAGTCAAGCTCTCTGAACCTTGAGCATGGCCCAAGCCCTGGACAAG 1332
QY 407 Ly---PheCysAlaValAlaGlnAlaAsnGlyLeuTyrProValAlaAlaAsnAsnArgAsnAlap 426
DB 1333 ACAGTCTTCCAGGCGGCAAGCTGATGGCTCTATCCCAATCCTCGGNAACGTCAGCT 1392
QY 426 heTrpHisCysValAsnGlyValThrTyrGlnGlnAsnCysGlnAlaGlyLeuValPheA 446
DB 1393 TCTACAGCTGTGACGCGGGGGCTGTTCCAGCAAGCTGCCGACAGGCTGGTGTTC 1452
QY 446 spThrSerCysAspCysCysAsnTrp 454
DB 1453 GCAACTCTCTGCAATGCTGCACCTGG 1478

RESULT 12

US-09-343-623-5
; Sequence 5, Application US/09343623
; Patent No. 6303118
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: A human chitinase, its recombinant
; TITLE OF INVENTION: production, its use for decomposing chitin, its use
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Hoffmann & Baron
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11758
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; SOFTWARE:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/343,623
; FILING DATE:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,839
; FILING DATE: 07-June-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Baron, Ronald J.
; REGISTRATION NUMBER: 29,281
; REFERENCE/DOCKET NUMBER: 294-26
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1713 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
US-09-343-623-5

Alignment Scores:

Pred. No.: 9,376-122 Length: 1713
Score: 1311.50 Matches: 244
Percent Similarity: 68.0% Conservative: 76
Best Local Similarity: 51.91% Mismatches: 131
Query Match: 52.33% Indels: 19
DB: 4 Gaps: 4

US-10-004-219B-14 (1-455) x US-09-343-623-5 (1-1713)

QY 2 GlnLeuThrCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPheMet 21
DB 79 AAACCTGGTCTGCTACTCTACCAACCTGGGCCAGTACAGACAGGAGGAGGCTCGCTTCCTG 138

QY 22 ProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArgGln 41
DB 139 CCCAAGGACTGGACCCACGCTTTGGACCCACCTCATCTAGCCCTTGGCTGGCATGACC 198

QY 42 AsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrGlnAlaPheAsnGly 61
DB 199 AACCAACGAGTGGACCACTGAGTGAATGACGAGACTCTCTACCAAGGAGTTCAATGGC 258

QY 62 LeuTyrAsnLysAsnSerGlnLeuTyrThrLeuAlaIleGlyClyTrpAsnPheGly 81
DB 259 CTGAAGAAGATGAATCCCAAGCTGAAGACCTGTGTAGCCATCGGAGGCTGGAAATTCGCG 318

QY 82 ThrAlaProPheThrAlaMetValSerThrProGluAsnAtgGlnThrPheIleThrSer 101
DB 319 ACTCAGAGTTTACAGATATGTTAGCCAGCGCAACACCTGTCAGACCTTTGTCAACTCG 378

QY 102 ValIleLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyrPro 121
DB 379 GCCATCAGGTTCTCGGCAAAATACAGCTTTGAGCGCTTGACCTTGACTGGGAGTACCCA 438

QY 122 GlySerArgGlySerProGlnAspLysHisLeuPheThrValLeuValGlnGluMet 141
DB 439 GGAAGCCAGGGGAGGCGCTCGCGTACAGAGGAGCGCTTCACAAACCTGGTACAGAGCTG 498

QY 142 ArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThrAla 161
DB 499 GCCAATGCTTCCAGCAGGAGCCAGACCTCAGCGGAAGAGCGCTTCTCTGAGTGCA 558

QY 162 AlaValAlaIleGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLysSerGln 181
DB 559 GCGGTTCAGCTGGGACAGACCTATGTGATGCTGATACGAGGTGGACAAAATGCCCCAG 618

QY 182 TyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluClyTyrThr 201
DB 619 AACCTGGATTTGTCAACCTTATGGCTTACGACTTCCATGGCTCTTGGGAGAAGTCCAG 678

QY 202 GlyGluAsnSerProLeuTyrLysTyrProThrAspThrGlySerAsnAlaTyrLeuAsn 221
DB 679 GGACATAACAGCCCTTACAGAGGCAAGAGAGTGTGGCAGCAGCCAGCTCAAC 738

QY 222 ValAspTyrValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeuIleVal 241
DB 739 GTGGATGCTGCTGCAACAGTGGCTGCAGAGGGGACCCCTGCCAGCAAGCTGATCCT 798

QY 242 GlyPheProThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIleGly 261
DB 799 GGCATGCTTACCTACGAGCGCTCTTCACCTGCGCTCTCATCAGACACCAAGAGTGGGG 858

QY 262 AlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyIleTrpAla 281
DB 859 GCCCAGCCAGCAGGTTGCGACTCAGGCCCTTACCAAGAAAGGAGGATGTGGCC 918

QY 282 TyrTrpGluIleCysThrPheLeuLysAsnGlyAlaThrGlnIleTrpAspAlaProGln 301
DB 919 TACTATGAAGTCTGCTCCTGG-----AAGGGGGCCACCAACAGAGATCCAGGATCAG 972

QY 302 GluValProTyrAlaTyrGlnGlyAsnValTrpValGlyTyrAspAsnIleLysSerPhe 321
DB 973 AAGGTGCCCTACATCTCTCGGGACAAACAGTGGTGGGCTTTGATGATGTGGAGAGCTTC 1032

QY 322 AspIleLysAlaGlnTrpLeuLysHisAsnLysPheGlyGlyAlaMetValTrpAlaIle 341
DB 1033 AAAACCAAGGTCAGCTATCTGAGCAGAGAGGAGCTGGGGGGCCCATGGTCTGGGCACTG 1092

QY 342 AspLeuAspAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSerThr 361
DB 1093 GACTTAGATGACTTTGGCGGCTTCTCTGCAACCAAGGCGCGATACCCCTCATCCAGAGC 1152

QY 362 LeuLysLysAlaLeuGlyLeuGlnSerAlaSerCysThrAlaProAlaGlnPro-IleGln 381
DB 1153 CTACGGCAGGAACGAATGGGTAAAGCCTTAACTGCCTGTCCATGTGAGGCCAGGTGT 1212

QY 381 uProlleThr-AlaAla-----ProSerGlySerGlyAsn- 392
DB 1213 GCCTGTGCACATGTGCTTCAGCTGTAGTCTTCCATATCTTGCCTTCAGGCCACCCAGAGC 1272

QY 393 -----GlySerGlySerSerSerSerGlyGlySerSerGlyGlySerG 407
DB 1273 TTGAAGTTTCAAAACCAAGTCAGCTCAGCCCTCTGAACCTGAGCATGGCCCAAGCCTGACAAG 1332

QY 407 ly---PheCysAlaValArgAlaAsnGlyLeuTyrProValAlaAsnAsnArgAsnAlaP 426
DB 1333 ACAGCTTCTGCCAGGGCAAGCTGATGGGCTCTATCCCAATCCTCGGGAACGGTCCAGCT 1392

QY 426 heTrpHisCysValAsnGlyValThrTyrGlnGlnAsnCysGlnAlaGlyLeuValPheA 446
DB 1393 TCTACAGCTGTGCAGCGGGGGGGCTGTTCAGCAAGAGTCCCGACAGCCCTGGTGTTC 1452

QY 446 spThrSerCysAspCysCysAsnTrp 454
DB 1453 GCAACTCTCTGCAATGTCTGCACCTGG 1478

RESULT 13

US-08-694-915-1
; Sequence 1, Application US/08694915
; Patent No. 5811535
; GENERAL INFORMATION:
; APPLICANT: Adamou, Julie
; APPLICANT: Kirkpatrick, Robert
; APPLICANT: Rosenberg, Martin
; TITLE OF INVENTION: HUMAN CARTILAGE GP39-LIKE GENE
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-2799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/694,915
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Han, William T
; REGISTRATION NUMBER: 34,344
; REFERENCE/DOCKET NUMBER: ATG50017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5219
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1433 base pairs

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
US-08-694-915-1

Alignment Scores:

Pred. No.: 9,88e-88 Length: 1433
Score: 970.50 Matches: 182
Percent Similarity: 69.11% Conservative: 73
Best Local Similarity: 49.32% Mismatches: 105
Query Match: 38.73% Indels: 9
DB: 1 Gaps: 5

US-10-004-219b-14 (1-455) x US-08-694-915-1 (1-1433)

QY 1 TyrGlnLeuThrCysThrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPhe 20
DB 115 TACAACTGGTTGCTTACTTACCACTGGTCCAGGACCGGAGAACCCAGGAAATTC 174
QY 21 MetProAspAsnIleAspProCysLeuLeuThrHisLeuIleTyrAlaPheAlaGlyArg 40
DB 175 ACCCTGAGATATGACCCCTCTCTATGCTCTCATCTCATCTATTCATTCATTCGACGATC 234
QY 41 GlnAsnAsnGluThrThrThrLeuGluTrpAsnAspValThrLeuTyrGlnAlaPheAsn 60
DB 235 GAAACCAACAAGGTATCATCAAGGACCAAGAGTGAAGTGTCTTACCAAGACCATCAAC 294
QY 61 GlyLeuLeuAsnLysAsnSerGlnLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPhe 80
DB 295 AGTCTCAAAACCAAGAAATCCAACTGAAATTCCTTGTCCATTCGAGGGTACCTGTTT 354
QY 81 GlyThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThr 100
DB 355 GGTTCAAAGGGTTCACCTATGTTGGATCTCTTACATCAGCTTGGAAATTCATTAAC 414
QY 101 SerValIleLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyr 120
DB 415 TCATAATCTGTTCTGAGGAACCAATACTTTGATGAGTGTGATGATGAAGTGGATCTAC 474
QY 121 ProGlySerArgGlySerProGlnAspLysPheHisLeuPheThrValLeuValGlnGlu 140
DB 475 CCAGATCAGAAA-----GAAACACTCAT----TTCACTGTGCTGATCATGAG 519
QY 141 MetArgGluAlaPheGluGlnGluAlaLysGlnLeuAsnLysProArgLeuMetValThr 160
DB 520 TTAGCAGAGGCTTTCAGAGGAGCTTCACAAATCCACCAAGAAAGCTTCTCTGACT 579
QY 161 AlaAlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSer 180
DB 580 GCGGGGTATCTCAGGGAGGCAAAATGATTAACAGCATCATCAAGTTCAGAAACTGGCA 639
QY 181 GlnTyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGly--- 199
DB 640 AAGATCTGGATTTCATCAACCTCTGTCTTGACTTCATGGGTCTTGGGAAAGGCC 699
QY 200 ---TyrThrGlyGluAsnSerProLeuTyrLysTyrProThrAspThrGlySerAsnAla 218
DB 700 CTTATCACTGGCCACACAGGCTTCTGACGAGGGGTGGCAGACAGAGGGCCCAAGCTCC 759
QY 219 TyrLeuAsnValAspTyrValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLys 238
DB 760 TACTAATCTGGAATGATGCTGGGGTACTGTGATACATAAGGGGAATGCCATCAGAGAAG 819
QY 239 LeuIleValGlyPheProThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThr 258
DB 820 GTGGTATGGGATCCCAATATGGGACTCTTTCACACTGGCCCTCT---CGAGAAACC 876
QY 259 GlyIleGlyAlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGly 278

DB 877 ACCGTGGGGGCCCTGCTCTGGCCCTGGAGCTGCTGGACCCATCACAGAGTCTTCAGGC 936
QY 279 IleTrpAlaTyrTyrGluIleCysThrPheLeuLysAsnGlyAlaThrGlnGlyTrpAsp 298
DB 937 TTCCTGGCCCTATTATGAGATCTGCCAGTTCTCTGAAA---GGAGCCAAAGATCACGGGGCTC 993
QY 299 AlaProGlnGluValProTyrAlaTyrGlnGlyAsnValTrpValGlyTyrAspAsnIle 318
DB 994 CAGGATCAGAGGTTCCCTACGAGTCAAGGGAACCACTGGGTGGCTATGATGATG 1053
QY 319 LysSerPheAspIleLysAlaGlnTrpLeuLysHisAsnLysPheGlyGlyAlaMetVal 338
DB 1054 AAGAGTATGGAGACCAAGGTTCAAGTCTTTAAAGATTTAAACCTGGAGAGCCATGATC 1113
QY 339 TrpAlaIleAspLeuAspAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeu 358
DB 1114 TGGTCTATTGACATGGATGACTTCACTGCAAAATCTTCGAACCAAGGGCCCTTACCTCTTT 1173
QY 359 IleSerThrLeuLysLysAlaLeuGly 367
DB 1174 GTCCAAGCAGTCAAGAGACCTTGGC 1200

RESULT 14

US-08-694-915-3

Sequence 3, Application US/08694915

Patent No. 5811535

GENERAL INFORMATION:

APPLICANT: Adamou, Julie

APPLICANT: Kirkpatrick, Robert

APPLICANT: Rosenberg, Martin

TITLE OF INVENTION: HUMAN CARTILAGE GP39-LIKE GENE

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:

ADDRESSEE: SmithKline Beecham Corporation

STREET: 709 Swedeland Road

CITY: King of Prussia

STATE: PA

COUNTRY: USA

ZIP: 19406-2799

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: Fast-Seq Version 1.5

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/694,915

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Han, William T

REGISTRATION NUMBER: 34,344

REFERENCE/DOCKET NUMBER: ATG50017

TELECOMMUNICATION INFORMATION:

TELEPHONE: 610-270-5219

TELEFAX: 610-270-5090

TELEX:

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 1526 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE:

ORIGINAL SOURCE:

US-08-694-915-3

[illegible]

1 TyrGlnLeuThrCysTyrPheThrAsnTrpAlaGlnTyrArgproGlyLeuGlyArgPhe 20
 135 TACAAACTGCTGTCTACTACACAGCTGGTCCCAGTACCGGAAGCGCATGGGAGCTGC 19
 21 MetProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArg 40

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: June 29, 2003, 23:54:39 ; Search time 165.913 Seconds
(without alignments)
4070.941 Million cell updates/sec

Title: US-10-004-219B-14
Perfect score: 2506
Sequence: 1 YQLTCYFTNWAQYRPLGRF.....QQNCOAGLVFTSDCCNCWA 455

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1055720 seqs, 742224136 residues

Total number of hits satisfying chosen parameters: 2111440

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -MODEL=frame+ p2n.model -DEV=xlsp
-Q=/cgn2_1/USPTO_spool/US10004219/runat_24062003_160230_8285/app_query.fasta_1.1877
-DB=Published Applications NA -QFMT=fastcap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blom62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10004219 @CNG 1 1 123 @runat_24062003_160230_8285
-NCPUS=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FCGPOP=6 -FCGPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

- 1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq*
- 2: /cgn2_6/prodata/1/pubpna/PCT_NEW_PUB.seq*
- 3: /cgn2_6/prodata/1/pubpna/US06_NEW_PUB.seq*
- 4: /cgn2_6/prodata/1/pubpna/US06_PUBCOMB.seq*
- 5: /cgn2_6/prodata/1/pubpna/US07_NEW_PUB.seq*
- 6: /cgn2_6/prodata/1/pubpna/PCTUS_PUBCOMB.seq*
- 7: /cgn2_6/prodata/1/pubpna/US08_NEW_PUB.seq*
- 8: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq*
- 9: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq*
- 10: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq*
- 11: /cgn2_6/prodata/1/pubpna/US10_NEW_PUB.seq*
- 12: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq*
- 13: /cgn2_6/prodata/1/pubpna/US60_NEW_PUB.seq*
- 14: /cgn2_6/prodata/1/pubpna/US60_PUBCOMB.seq*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	2506	100.0	1625	US-10-004-219B-2
2	2089.5	83.4	1525	US-10-004-219B-3
3	1251	49.9	1038	US-10-202-436A-14
4	985.5	39.3	1925	US-10-097-340-44

5	957.5	38.2	1681	10	US-09-215-077A-4	Sequence 4, Appli
6	957.5	38.2	1681	10	US-09-262-213A-4	Sequence 4, Appli
7	923.5	36.9	1391	10	US-09-822-830A-402	Sequence 402, App
8	873.5	34.9	1474	10	US-09-765-231A-8	Sequence 8, Appli
9	711.5	28.4	1470	9	US-10-218-743-40	Sequence 40, Appli
c 10	711.5	28.4	1470	9	US-10-218-743-42	Sequence 42, Appli
c 11	711.5	28.4	1527	9	US-10-218-743-37	Sequence 37, Appli
c 12	711.5	28.4	1527	9	US-10-218-743-39	Sequence 39, Appli
c 13	711.5	28.4	1621	9	US-10-218-743-34	Sequence 34, Appli
c 14	711.5	28.4	1621	9	US-10-218-743-36	Sequence 36, Appli
c 15	708.5	28.3	1608	9	US-10-218-743-20	Sequence 20, Appli
c 16	708.5	28.3	1608	9	US-10-218-743-22	Sequence 22, Appli
c 17	708.5	28.3	1665	9	US-10-218-743-17	Sequence 17, Appli
c 18	708.5	28.3	1665	9	US-10-218-743-19	Sequence 19, Appli
c 19	708.5	28.3	1752	9	US-10-218-743-14	Sequence 14, Appli
c 20	708.5	28.3	1752	9	US-10-218-743-16	Sequence 16, Appli
c 21	646	25.8	449	10	US-09-960-352-678	Sequence 678, App
c 22	472	18.8	1271	9	US-09-923-844B-1	Sequence 1, Appli
c 23	444.5	17.7	1140	9	US-09-938-842A-1684	Sequence 1684, App
c 24	444	17.7	1797	10	US-09-974-300-646	Sequence 646, App
c 25	380.5	15.2	410	10	US-09-960-352-7364	Sequence 7364, Ap
c 26	380.5	15.2	414	10	US-09-960-352-13050	Sequence 13050, A
c 27	380.5	15.2	415	10	US-09-960-352-2589	Sequence 2589, Ap
c 28	380.5	15.2	424	10	US-09-960-352-3057	Sequence 3057, Ap
c 29	378.5	15.1	418	10	US-09-960-352-3072	Sequence 3072, Ap
c 30	363	14.5	399	10	US-09-960-352-265	Sequence 265, App
c 31	361.5	14.4	411	10	US-09-960-352-14919	Sequence 14919, A
c 32	352	14.0	398	10	US-09-960-352-14460	Sequence 14460, A
c 33	347.5	13.9	384	10	US-09-960-352-5721	Sequence 5721, Ap
c 34	342	13.6	427	10	US-09-960-352-3315	Sequence 3315, Ap
c 35	341	13.6	438	10	US-09-960-352-1882	Sequence 1882, Ap
c 36	339	13.5	422	10	US-09-960-352-5957	Sequence 5957, Ap
c 37	339	13.5	423	10	US-09-960-352-12100	Sequence 12100, A
c 38	338	13.5	430	10	US-09-960-352-1189	Sequence 1189, Ap
c 39	336	13.4	419	10	US-09-960-352-10172	Sequence 10172, A
c 40	336	13.4	427	10	US-09-960-352-4978	Sequence 4978, Ap
c 41	335	13.4	426	10	US-09-960-352-4979	Sequence 4979, Ap
c 42	334	13.3	419	10	US-09-960-352-4053	Sequence 4053, Ap
c 43	334	13.3	422	10	US-09-960-352-10642	Sequence 10642, A
c 44	332	13.2	410	10	US-09-960-352-8317	Sequence 8317, Ap
c 45	331	13.2	429	10	US-09-960-352-14230	Sequence 14230, A

ALIGNMENTS

RESULT 1
US-10-004-219B-2
; Sequence 2, Application US/10004219B
; Publication No. US20030087414A1
; GENERAL INFORMATION:
; APPLICANT: Macrozyme
; APPLICANT: Aerts, Johannes M.F.G.
; APPLICANT: Boot, Rolf G.
; TITLE OF INVENTION: A mammalian mucinase, its recombinant production, and
; TITLE OF INVENTION: its use in therapy or prophylaxis against diseases in
; TITLE OF INVENTION: which mucus is involved or infection diseases
; FILE REFERENCE: 2183-5136US
; CURRENT APPLICATION NUMBER: US/10/004,219B
; CURRENT FILING DATE: 2001-11-02
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1625
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Combined DNA/RNA Molecule: human
; OTHER INFORMATION: AMCase cDNA sequence and deduced amino acid
; OTHER INFORMATION: sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: human AMCase
; OTHER INFORMATION: cDNA sequence and deduced amino acid sequence
; FEATURE:

NAME/KEY: CDS
LOCATION: (104)...(1531)
US-10-004-219B-2

Alignment Scores:

Pred. No.: 3,54e-280 Length: 1625
Score: 2506.00 Matches: 455
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-004-219B-14 (1-455) x US-10-004-219B-2 (1-1625)

QY	1	TyrGlnLeuThrCysTyrPheThrAsnTyrAlaGlnTyrArgProGlyLeuGlyArgPhe	20
DB	167	TACCACTGACATGCTTACCACTGAGCCAGTACCGCCAGGCTTGGGGCGCTTC	226
QY	21	MetProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArg	40
DB	227	ATGCCTGACAACATCGACCCCTGCTGTACCCACCTGATCTACGCTTTGCTGGGAGG	286
QY	41	GlnAsnAsnGluIleThrThrIleGluTyrAsnAspValThrLeuTyrGlnAlaPheAsn	60
DB	287	CAGAACACGAGATCACCACTCGAATCGAATGGAACGATGTGACTCTTACCAAGCTTTCAAT	346
QY	61	GlyLeuLeuAsnIleAsnSerGlnLeuLeuThrLeuAlaIleGlyGlyTyrAsnPhe	80
DB	347	GGCTTGAAAANTGAACACGCGAGCTGAAAACCTCTCTGGCCATTTGGAGCTGGAACCTTC	406
QY	81	GlyThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThr	100
DB	407	GGGACTGCCCTTTTCACTGCCATGGTTTCTACTCTCGAGAACCGCCAGACTTTTCATCACC	466
QY	101	SerValIleLeuPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyr	120
DB	467	TCAGTCATCAAAATTCCTGGCCAGTATGATTTGACGGGCTGGACTTTGACTGGGAGTAC	526
QY	121	ProGlySerArgGlySerProGlnAspLeuAspLeuPheThrValLeuValGlnGlu	140
DB	527	CTGGCTCTCGTGGGAGCCCTCTCAGGACAGCATCTTTCATCTGCTCTGGTGGAGGAA	586
QY	141	MetArgGluAlaPheGluGlnGluAlaLeuGlnIleAsnLysProArgLeuMetValThr	160
DB	587	ATGCGTGAAGCTTTTGAGCAGGAGGCCAAGCAGATCAACAGCCAGGCTGATGGTCACT	646
QY	161	AlaAlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSer	180
DB	647	GCTGCAGTAGCTGCTGGCATCTCCAAATATCCAGTCTGGCTATGAGATCCCAACTGTCA	706
QY	181	GlnTyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr	200
DB	707	CAGTACTCGACTACATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT	766
QY	201	ThrGlyGluAsnSerProLeuTyrLysTyrProThrAspThrGlySerAsnAlaTyrLeu	220
DB	767	ACTGGAGAGACAGCCCTCTACAAATACCCGACTGACACCGCCAGCAGCAAGCTTACTC	826
QY	221	AsnValAspTyrValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeuIle	240
DB	827	AATGTGGATTATGTCATGAACCTACTCGAAGGACAATGGAGCACCAGCTGAGAAGCTCATC	886
QY	241	ValGlyPheProThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIle	260
DB	887	GTGGATTCCCTACCTATGACACAACTTCATCTGAGCAACCCCTCCCAACTGGAAT	946
QY	261	GlyAlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyIleTrp	280
DB	947	GGTGGCCCCACCTCTGGTGTCTGCTGGGCCCTATGCCAAGAGCTGGATCTGG	1006
QY	281	AlaTyrTyrGluIleCysThrPheLeuLysAsnGlyAlaThrGlnGlyTrpAspAlaPro	300
DB	1007	GCTTACTACGAGATCTGTACCTTCTGAAAAATGGAGCCACTCAGGATGGGATGCCCT	1066

QY	301	GlnGluValProTyrAlaTyrGlnGlyAsnValTrpValGlyTyrAspAsnIleLysSer	320
DB	1067	CAGGAAGTGCCTTATGCTTATCAGGGCAATGTGTGGTGGCTATGACACATCAAGAGC	1126
QY	321	PheAspIleLysAlaGlnTrpLeuLysHisAsnLysPheGlyGlyAlaMetValTrpAla	340
DB	1127	TTTCGATATTAAAGCTCAATGGCTTAAGCACAAATAATTTGGAGGCGCATGCTCTGGGCG	1186
QY	341	IleAspLeuAspAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSer	360
DB	1187	ATTGATCTGGATGACTTCACTGGCACTTTCTGCAACACAGGGCAAGTTTCCCTAATCTCC	1246
QY	361	ThrLeuLysLysAlaLeuGlyLeuGlnSerAlaSerCysThrAlaProAlaGlnProIle	380
DB	1247	ACCTGAGAAGAGCCCTCGGCTTCAGAGTGCAAGTTCACGGCTCCAGCTCAGGCCATT	1306
QY	381	GluProIleThrAlaAlaProSerGlySerGlyAsnGlySerGlySerSerSerGly	400
DB	1307	GAGCCAATAACTGTCTCCAGTGGCAGCGGGAACGGGAGCGGAGTAGCAGCTCTGGA	1366
QY	401	GlySerSerGlyGlySerGlyPheCysAlaValAlaArgAlaAsnGlyLeuTyrProValAla	420
DB	1367	GGCAGCTCGGGAGGCGAGTGGATTCTGTCTGTGAGAGCCAAAGCGCTCTACCCCGTGGA	1426
QY	421	AsnAsnArgAsnAlaPheTrpHisCysValAsnGlyValThrTyrGlnAsnCysGln	440
DB	1427	ATAACAGAAATGCCCTTTGGCACTCGGTGATGATGATGATGATGATGATGATGATGAT	1486
QY	441	AlaGlyLeuValPheAspThrSerCysAspCysCysAsnTrpAla	455
DB	1487	GCGGGCTGTCTTCGACACAGCTGTGATTGCTGCAACTGGGCA	1531

RESULT 2

US-10-004-219B-3
Sequence 3, Application US/10004219B
Publication No. US2003008741A1

GENERAL INFORMATION:

APPLICANT: Macrozyme
APPLICANT: Aerts, Johannes M.F.G.
APPLICANT: Boot, Rolf G.
TITLE OF INVENTION: A mammalian mucinase, its recombinant production, and its use in therapy or prophylaxis against diseases in which mucus is involved or infection diseases
TITLE OF INVENTION: its use in therapy or prophylaxis against diseases in which mucus is involved or infection diseases
FILE REFERENCE: 2183-5136US
CURRENT APPLICATION NUMBER: US/10/004,219B
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3

LENGTH: 1525

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE: Description of Combined DNA/RNA Molecule: mouse

OTHER INFORMATION: AMCase cDNA sequence and deduced amino acid

OTHER INFORMATION: sequence

FEATURE: Description of Artificial Sequence: mouse AMCase

OTHER INFORMATION: cDNA sequence and deduced amino acid sequence

FEATURE:

NAME/KEY: CDS

LOCATION: (1)...(1419)

US-10-004-219B-3

Alignment Scores:

Pred. No.: 5,79e-232 Length: 1525
Score: 2089.50 Matches: 369
Percent Similarity: 90.05% Conservative: 40
Best Local Similarity: 81.28% Mismatches: 42
Query Match: 83.38% Indels: 3
DB: 9 Gaps: 1

US-10-004-219B-14 (1-455) x US-10-004-219B-3 (1-1525)

QY 1 TyrGlnLeuThrCysTyrPheThrAsnTyrAlaGlnTyrArgProGlyLeuGlyArgPhe 20
DB 64 TACATCTGATGATGCTATTTTCCACCACTGGGCCAGATATCGGCCAGGCTCTGGGAGCTTC 123
QY 21 MetProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArg 40
DB 124 AAGCCTGATGACATTAACCCCTGCTGTGTAATCTACCTGATCTATGCTTGTGGATG 183
QY 41 GlnAsnAsnGluIleThrThrIleGluTyrAsnAspValThrLeuTyrGlnAlaPheAsn 60
DB 184 CAGAACATGAGATCACCACCATAGATGAATGATGTTACTCTCTATAAGCTTTCAAT 243
QY 61 GlyLeuLeuAsnLysAsnSerGlnLeuLysThrLeuLeuAlaIleGlyGlyTyrAsnPhe 80
DB 244 GACTTGAAAAACAGGACAGCAACTGAAACCCCTCTGGCAATTGGAGCTGGAACTTT 303
QY 81 GlyThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThr 100
DB 304 GGAACCTGCTCTTCACTACCATGGTTTCCACTTCTCAGAACCGCCAGACCTTCATTACC 363
QY 101 SerValIleLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyr 120
DB 364 TCAGTCATCAAAATTTCTGCGTCAGTATGGTTTGTATGGACTGGACCTGGAGTAATAC 423
QY 121 ProGlySerArgGlySerProGlnAspLysHisLeuPheThrValLeuValGlnGlu 140
DB 424 CCAGCTCAGTGGAGCCCTCTCAGCAAGCATCTTCTACTGTCTCTGGTAGGAA 483
QY 141 MetArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThr 160
DB 484 ATGCGTGAAGCTTTTTCAGCAGGAGGCTATTGAGAGCAACAGCCAGACTGATGGTTACT 543
QY 161 AlaAlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSer 180
DB 544 GCTGCTGAGCTGGTGGGATTTTCCAACTCCAGGCTGGCTATGAGATCCCTGAACTTCT 603
QY 181 GlnTyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr 200
DB 604 AAGTACCTGGATTTTATCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 663
QY 201 ThrGlyGluAsnSerProLeuTyrIleTyrProThrAspThrGlySerAsnAlaTyrLeu 220
DB 664 ACTGGGGAGATAGTCTCTTTTACAAATACCTCTACTGAGACTGGTAGCAATGCCTACCTC 723
QY 221 AsnValAspTyrValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeuIle 240
DB 724 AATGTGGATTATGTATGAACATTTTGGAAAGAACAAATGGAGCCCGAGCTGAGAAGCTCAT 783
QY 241 ValGlyPheProThrTyrGlyHisPheIleLeuSerAsnProSerAsnThrGlyIle 260
DB 784 GTTGGATTCCCAAGATATGACACACTTCTATCTCTGAGAAACCCCTCTGATTAATGAAT 843
QY 261 GlyAlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyIleTrp 280
DB 844 GGTGCCCTTACCTCTGTGTATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 903
QY 281 AlaTyrTyrGluIleCysThrPheLeuLysAsnGlyAlaThrGlnGlyTrpAspAlaPro 300
DB 904 GCCTACTATGAGATTTGCACTTCTTGAAGATGGAGCCACTGAGCTCTGGGATGCTCTCC 963
QY 301 GlnGluValProTyrAlaTyrGlnGlyAsnValTrpValGlyTyrAspAsnIleLysSer 320
DB 964 CAAAGATGCTCTATGCCCTATAGGCCAACGAGTGGCTTGGCTGTATGACAAATATCAAGAGC 1023
QY 321 PheAspIleLysAlaGlnTrpLeuLysHisAsnLysPheGlyGlyAlaMetValTrpAla 340
DB 1024 TTCAGTGTAAAGCTCAGTGGCTTAGCAGACAAATTTTGGAGGTGCCATGATCTGGGCC 1083
QY 341 IleAspLeuAspAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSer 360
DB 1084 ATTGACCTTGTGATGACTTCACTGGCTCTTCTGTGATCAGGAGAAATTTCTCTGACTTCT 1143

QY 361 ThrLeuLysLysAlaLeuGlyLeuGlnSerAlaSerCysThrAlaProAlaGlnProIle 380
DB 1144 ACTTTGAACAAGCCCTTGGCATATCCACTGAGGTTGCACAGCTCTCTGACGTGCTTCC 1203
QY 381 GluProIleThrAlaAlaProSerGlySerGlyAsnGlySerGlySerSerSerSerGly 400
DB 1204 GAGCCAGTGACTACTCTCTCCA-----GGAAGTGGAGTGGGGGTGGAAGCTCCGGA 1254
QY 401 GlySerSerGlyGlySerGlyPheCysAlaValAlaArgAlaAsnGlyLeuTyrProValAla 420
DB 1255 GGAAGCTCTGGAGGAGTGGATTCTGTGCCGAAGCAGATGGCTCTTACCTCTGTGGCA 1314
QY 421 AsnAsnArgAsnAlaPheTrpHisCysValAsnGlyValThrTyrGlnGlnAsnCysGln 440
DB 1315 GATGACAGAAATGCTTTTGGCAGTGCATCAATGAATCACATACAGCAGCATTTGTCAA 1374
QY 441 AlaGlyLeuValPheAspThrSerCysAspCysCysAsnTrp 454
DB 1375 GCAGGCTTGTTTTGTATACAGCTGTAATTTGCTGCAACTGG 1416

RESULT 3

US-10-202-436A-14
; Sequence 14, Application US/10202436A
; Publication No. US20030049261A1
; GENERAL INFORMATION:
; APPLICANT: ELIAS, JACK A
; APPLICANT: ZHU, ZHOU
; TITLE OF INVENTION: METHODS COMPOSITIONS AND KITS RELATING TO CHITINASES AND CHITINASE
; FILE REFERENCE: 044574-5107
; CURRENT APPLICATION NUMBER: US/10/202,436A
; CURRENT FILING DATE: 2002-10-17
; PRIOR APPLICATION NUMBER: U.S. 60/307,432
; PRIOR FILING DATE: 2001-07-24
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 1038
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: AMCase sense in situ hybridization probe
US-10-202-436A-14

Alignment Scores:
Pred. No.: 4,528-135 Length: 1038
Score: 1251.00 Matches: 237
Percent Similarity: 92.69% Conservative: 4
Best Local Similarity: 91.15% Mismatches: 13
Query Match: 49.92% Indels: 6
DB: 9 Gaps: 2

US-10-004-219B-14 (1-455) x US-10-202-436A-14 (1-1038)

QY 202 GlyGluAsnSerProLeuTyrLysTyrProThrAspThrGlySerAsnAlaTyrLeuAsn 221
DB 2 GGAGAGACAGAGCCCTCTTACAAATACCGAGTGCACCGGAGCAACGCTTACCTCAAT 61
QY 222 ValAspTyrValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeuIleVal 241
DB 62 GTGGATTATGTCATGAACCTACTTGAAGGACAAATGAGCAGCAGCTGAGAAGCTCATCGTT 121
QY 242 GlyPheProThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIleGly 261
DB 122 GGATTCCTTACCTATGGACACAACTTCATCTCTGAGCAACCCCTCCAAACTGGAAATGGT 181
QY 262 AlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyIleTyrAla 281
DB 182 GCCCCACCTCTGGTGTGCTGCTCTGCTGGGCCCTTATGCCAGAGTCTGGGATCTGGGCT 241
QY 282 TyrTyrGluIleCysThrPheLeuLysAsnGlyAlaThrGlnGlyTrpAspAlaProGln 301

```
Db 242 TACTACGAGCTGTACCTCTCTGAAAAATGGAGCCACTCAGGATGGGATGCCCTCAG 301
Qy 302 GluValProTyrAlaTyrGlnGlyAsnValTyrValGlyTyrAspAsnIleYsSerPhe 321
Db 302 GAAGTCCCTTATGCCCTATCAGGCAATGTGTGGTTGGCTATGACACAGCTCAGAGCTTC 361
Qy 322 AspIleYsAlaGlnTrpLeuYsHisAsnLysPheGlyGlyAlaMetValTrpAlaIle 341
Db 362 GATATTAAAGGCTCAATGGCTTAAAGCACAAATCTGGAGGCGCCATGTGTGGGCCATT 421
Qy 342 AspLeuAspAspPheThrGlyThrPheCysAsnGlnGlyPheProLeuIleSerThr 361
Db 422 GATCTCGATGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 481
Qy 362 LeuLysAlaLeuGlyLeuGlnSerAlaSerCysThrAlaProAlaGlnProIleGlu 381
Db 482 CTGAAGAAGGCCCTCGGGCTGCAAGATGCAATGTTCAGGCTCCAGCTCAGCCCATITAG 541
Qy 382 ProIleThrAlaAlaProSerGlySerGlyAsnGlySerGlySerSerSerGlyGly 401
Db 542 CCAATTAAGTCTGCTCCAGTGGCAGCGGCAACGGGAGGGAGTAGCAGCTCTGGAGGC 601
Qy 402 SerSerGlyGlySerGlyPhe----CysAlaValArgAlaAsnGlyLeuTyrProValAla 420
Db 602 AGCTCGGAGGCGAGTGGATCTTGTCTGGCAGAGCAAAACGAGCTCTAACCCCGTGGGC 661
Qy 421 Asn---AsnArgAsnAlaPheTrp-HisCysValAsnGlyValThrTyrGln-GlnAsn- 438
Db 662 AAATTACCAAGAGATGCTTCTGGCAGCTGCGTGAATGGATGATGATGATGATGATGATG 721
Qy 439 CysGlnAlaGlyLeuVal-PheAspThrSerCysAspCysCysAsnTrpAla 455
Db 722 TGCCAGGCGCGGCTGTCTCTTCGAGACAGCTGTGTAATGCTGCAACTGGGCA 773

RESULT 4
US-10-097-340-44
; Sequence 44, Application US/10097340
; Publication No. US20030087250A1
; GENERAL INFORMATION:
; APPLICANT: John MONAHAN
; APPLICANT: Manjula GANNAVARAPU
; APPLICANT: Sebastian HOERSCH
; APPLICANT: Shubhangi KAMATKAR
; APPLICANT: Steve G. KOVATS
; APPLICANT: Rachel E. MEYERS
; APPLICANT: Michael MORRISSEY
; APPLICANT: Peter OLANDT
; APPLICANT: Ami SEN
; APPLICANT: Peter VEIBY
; APPLICANT: Gordon B. MILLS
; APPLICANT: Robert C. BAST, Jr.
; APPLICANT: Karen LU
; APPLICANT: Rosemarie SCHMANDT
; APPLICANT: Xumei ZHAO
; APPLICANT: Karen GLATT
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
; TITLE OF INVENTION: Assessment, Prevention, and Therapy of Ovarian Cancer
; FILE REFERENCE: MRI-030
; CURRENT APPLICATION NUMBER: US/10/097,340
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/276,026
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/323,580
```

```
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 1925
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-097-340-44

Alignment Scores:
Pred. No.: 6,748-104 Length: 1925
Score: 985.50 Matches: 195
Percent Similarity: 65.74% Conservative: 64
Best Local Similarity: 49.44% Mismatches: 120
Query Match: 39.33% Indels: 15
DB: 7 Gaps: 7

US-10-004-219B-14 (1-455) x US-10-097-340-44 (1-1925)

Qy 1 TyrGlnLeuThrCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPhe 20
Db 190 TACAACTGGTCTGCTACTACTACACAGCTGGTCCAGTACCCGGAAGGATGGAGCTGC 249
Qy 21 MetProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArg 40
Db 250 TTCCAGAGATGCCCTTGACCGCTTCTCTGTATCCACATCATCAGCTTTGCCAATATA 309
Qy 41 GlnAsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTrpGlnAlaPheAsn 60
Db 310 AGCAACGATCATCATCGACACCTGGGAGTGAATGATGACGCTCTACGGCATGCTCAAC 369
Qy 61 GlyLeuLysAsnLysAsnSerGlnLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPhe 80
Db 370 ACACCTCAAGAACAGGACCCCAACCTGAAGACTCTCTTGTCTGTCGGAGATGGAATTT 429
Qy 81 GlyThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThr 100
Db 430 GGGTCTCAAGATTTTCCAAAGATAGCTCCACACCCAGAGTCCGCGGACTTTTCATCAAG 489
Qy 101 SerValIleLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyr 120
Db 490 TCAGTACCGCCCATTCCTGGCCACCCATGCTTGTATGGCTGGACCTTCGCTGGCTCTAC 549
Qy 121 ProGlySerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValGlnGlu 140
Db 550 CTGGAGCGGAGA-----GACAAACAGCATTTTACCACCCTAATCAAGGAA 594
Qy 141 MetArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThr 160
Db 595 ATGAAGCGCGAATTTATAAAGGAAGCC---CAGCAGGAGGAAAGCAGCTCTGCTCAGC 651
Qy 161 AlaAlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSer 180
Db 652 GCAGCAGCTGTCTGGCGGGAAGGTACCATTTGACAGCTATGACATTCAGGATATCC 711
Qy 181 GlnTyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr 200
Db 712 CAACACCTGGATTTTCATTAGCATCATGACCTACGATTTTCATGAGCTGGCGTGGGACC 771
Qy 201 ThrGlyGluAsnSerProLeuTyrLysTyrProThrAspThrGlySerAsnAlaTyrLeu 220
Db 772 ACAGGCCATCATCAGTCCCTCTGTCGAGGTACAGGATGCAAGTCCCTGACAGATTCAGC 831
Qy 221 AsnValAspTyrValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGlyLeuIle 240
Db 832 AACACTGACTATGCTGTGGGTACATGTTGAGGCTGGGGGCTCTCCGCCAGTAACTGGTG 891
Qy 241 ValGlyPheProThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIle 260
Db 892 ATGGGCATCCCACTTCGGGAGGAGCTTCACTCTGGCTTCT--TCTGAGACTGGTGT 948
Qy 261 GlyAlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyIleTrp 280
```


Sequence 4, Application US/09262213A
Patent No. US20020090658A1

GENERAL INFORMATION:

APPLICANT: PRICE, PAUL
APPLICANT: JOHANSEN, JULIA
TITLE OF INVENTION: ASSAY FOR YKL-40 AS A MARKER FOR DEGRADATION OF
FILE OF INVENTION: MAMMALIAN CONNECTIVE TISSUE MATRICES
FILE REFERENCE: 4077-895412US
CURRENT APPLICATION NUMBER: US/09/262,213A
CURRENT FILING DATE: 1999-03-04
PRIOR APPLICATION NUMBER: US 08/089,989
PRIOR FILING DATE: 1993-07-09
PRIOR APPLICATION NUMBER: PCT/US94/07754
PRIOR FILING DATE: 1994-07-08
PRIOR APPLICATION NUMBER: US 08/581,527
PRIOR FILING DATE: 1996-04-17
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn version 3.0
SEQ ID NO 4

LENGTH: 1681

TYPE: DNA

ORGANISM: Homo sapiens

US-09-262-213A-4

Alignment Scores:

Pred. No.:	9,61e-101	Length:	1681
Score:	957.50	Matches:	194
Percent Similarity:	65.56%	Conservative:	63
Best Local Similarity:	49.49%	Mismatches:	123
Query Match:	38.21%	Indels:	13
DB:	10	Gaps:	6

US-10-004-219B-14 (1-455) x US-09-262-213A-4 (1-1681)

QY	1	TyrGlnLeuThrCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPhe	20
DB	135	TACAAACTGGTCTGCTACTACTACCACTGGTGTGGTACCGGAGCGATGGAGCTGC	194
QY	21	MetProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArg	40
DB	195	TTCCAGATGCCCTTGACCGCTCTCTGTGTACCCACATCATCTACAGCTTGGCAATATA	254
QY	41	GlnAspAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrGlnAlaPheAsn	60
DB	255	AGCAACGATCATCATCGACACCTGGGAGTGGAATGATGACGCTTACCGCATGTCTCAAC	314
QY	61	GlyLeuIleAsnLysAsnSerGlnLeuIleThrLeuAlaIleGlyGlyTrpAsnPhe	80
DB	315	ACACTCAACAACAACGAACCCCAACCTGGAAGACTCTCTGTGTGGAGGATGGAACTTT	374
QY	81	GlyThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThr	100
DB	375	GGGTCTCAAGATTTTCCAAAGTAGCTTCAACACCCAGAGTCCCGGACTTTCATCAAG	434
QY	101	SerValIleLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyr	120
DB	435	TCAGTACCGCCATTTCTGCGCACCATGCTTTGATGGCGTGAACCTTGCCTGCTCTAC	494
QY	121	ProGlySerArgGlySerProProGlnAspLysPheHisLeuPheThrValLeuValGlnGlu	140
DB	495	CCTGGACGCGA-----GACAAACACCATTTTACCAACCTTAATCAAGAA	539
QY	141	MetArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThr	160
DB	540	ATGAAGGCCGAATTTAATAGGAGCC---CAGCCAGGGAAGAGCAGCTCTGCTCAGC	596
QY	161	AlaAlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSer	180
DB	597	GCAGCACTGTCTCGGGAAGGTCAACATTGACAGCAGCTATGACATTGCCAAGATATCC	656
QY	181	GlnTyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr	200
DB	657	CAACACCTGGATTTTACATCATCATGACCTACGATTTTTCATGGCCCTCGGCTGGGACC	716

QY	201	ThrGlyGluAsnSerProLeuTyrIleTyrPheThrAspThrGlySerAsnAlaTyrLeu	220
DB	717	ACAGGCCATCAGCTCCCTCAGGCGAGGTCAAGGATGCAAGTCTCAGACATTCAGC	776
QY	221	AsnValAspTyrValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeu	240
DB	777	AACACTGACTATGCTGGGTGACATGTTGAGGCTGGGGCTCTCCGCCAGTAGTGGT	836
QY	241	ValGlyPheProThrTyrHisAsnPheIleLeuSerAsnProSerAsnThrGlyIle	260
DB	837	ATGGGCATCCCACTTCGGAGGAGCTTCACTCTGGCTTCT---TCTGAGACTGGTGT	893
QY	261	GlyAlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyIleTrp	280
DB	894	CCAGCGCCAATCTCAGGACCGGGAATTCAGGCCGTTTCAACCAAGGAGGAGGCCCTT	953
QY	281	AlaTyrTyrGluIleCysThrPheLeuLysAsnGlyAlaThrGlnGlyTrpAspAlaPro	300
DB	954	GCCTACTATGAGATCTGTGACTTCTCTCCGC---GGAGCCACAGTCCATAGAACCTCGGC	1010
QY	301	GlnGluValProTyrAlaTyrGlnGlyAsnValTrpValGlyTyrAspAsnIleLysSer	320
DB	1011	CAGCAGGTCCTTATGCCACCAAGGCAACCGTGGTAGGATACACACAGCAGGAAGC	1070
QY	321	PheAspIleLysAlaGlnTrpLeuLysHisAsnLysPheGlyGlyAlaMetValTrpAla	340
DB	1071	GTCAAAAGCAAGTGCAGTACCTGAAGGATAGGACGAGTGGCAGCGCCATGTATGGGCC	1130
QY	341	IleAspLeuAspAspPheThrGlyThrPheCysAsnGlnGly---LysPheProLeu	359
DB	1131	CTGGACCTGGATGACTTCCAGGCTCTCTTCTCGCCAGGATCTCGCTTCCCTCTCACC	1190
QY	360	SerThrLeuLysLysAlaLeuGlyLeuGlnSerAlaSerCysThrAlaProAlaGlnPro	379
DB	1191	ATGGCATCAGGATGACCTGCTGC---ACGTAGGCCCTCTGTCTCTGCACACAGCAGGGG	1249
QY	380	IleGluProIleThrAlaAlaProSerGlySerGly	391
DB	1250	GCCAAGGAT-----GCCCGTCCCGCTCTGCGC	1276

RESULT 7

US-09-822-830A-402
Sequence 402, Application US/09822830A
Patent No. US20020142952A1
GENERAL INFORMATION:
APPLICANT: Genetics Institute, Inc.
APPLICANT: Wong, Gordon G.
APPLICANT: Clark, Hilary
APPLICANT: Fechtel, Kim
APPLICANT: Agostino, Michael J.
APPLICANT: Howes, Steven H.
APPLICANT: Resnick, Richard J.
APPLICANT: Gulukota, Kamalakara
APPLICANT: Graham, James R.
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
FILE REFERENCE: GIN 6402
CURRENT APPLICATION NUMBER: US/09/822,830A
CURRENT FILING DATE: 2001-03-29
PRIOR APPLICATION NUMBER: 60/195,604
PRIOR FILING DATE: 2000-04-06
NUMBER OF SEQ ID NOS: 631
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 402
LENGTH: 1391
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: 1353_
OTHER INFORMATION: n-a, c, g, or t
US-09-822-830A-402

Alignment Scores:

Pred. No.: 6,24e-97 Length: 1391
Score: 923.50 Matches: 176
Percent Similarity: 67.48% Conservative: 73
Best Local Similarity: 47.70% Mismatches: 111
Query Match: 36.85% Indels: 9
DB: 10 Gaps: 5

US-10-004-219B-14 (1-455) x US-09-822-830A-402 (1-1391)

```
QY 1 TyrGlnLeuThrCysTyPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPhe 20
DB 76 TACAAACTGGTTGGGACTTTTACCAACTGGTCCAGGACCGCAGAAACCGAGAAATTC 135

QY 21 MetProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArg 40
DB 136 ACCCTGAGATATGACCCCTTCTATGCTCTCATCTCATCTATTCATTCGCCAGCATC 195

QY 41 GlnAsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrGlnAlaPheAsn 60
DB 196 GAAACCAACAAGGTATCATCAAGGACAAGAGTGAAGTGTCTCTACCAACCATCAAC 255

QY 61 GlyLeuLysAsnLysAsnSerGlnLeuLysThrLeuLeuAlaIleGlyGlyTrpAsnPhe 80
DB 256 AGTCTCAAAACCAAGAAATCCCAAACTGAAATTCCTTGTCCATTGGAGGGTACCTGTT 315

QY 81 GlyThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThr 100
DB 316 GGTTCAAAGGTTTCACCTATGTTGGATCTCTTACATCAACGCTTGAATTCATTAAC 375

QY 101 SerValIleLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyr 120
DB 376 TCCATAATCTGTTCTGAGGAACCAATACTTTGATGAGTGTGATGAAGCTGGATCTAC 435

QY 121 ProGlySerArgGlySerProGlnAspLysHisLeuPheThrValLeuValGlnGlu 140
DB 436 CCAGATCAGAAA-----GAAACACTCAT---TTCACCTGTGCTGATTCATGAG 480

QY 141 MetArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThr 160
DB 481 TTAGCAGAAGCTTTTCAGAAAGACTTCACAAATCCACCAGGAAGGCTTCTTGTGACT 540

QY 161 AlaAlaValAlaAlaGlyLysSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSer 180
DB 541 GCGGGCGTATCTGAGGAGGCAATGATTGATAACAGCTATCAAGTTTGAGAAACTGGCA 600

QY 181 GlnTyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGly--- 199
DB 601 AAAGATCTGGATTTTCATCAACTCTGCTGCTTGAATTCCTGAGTCTTGGGAAAGCCC 660

QY 200 ---TyrThrGlyGluAsnSerProLeuTyrLysTyrProThrAspThrGlySerAsnAla 218
DB 661 CTTATCATTGGCCACACAGCCCTCTGACGAGGGGTGCGACAGACAGAGGGCCAGCTCC 720

QY 219 TyrLeuAsnValAspTyrValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLys 238
DB 721 TACTACAATGTGGAATATGCTGGGGTACTGGATACATAAGGAATGCCATCAGAGAAG 780

QY 239 LeuIleValGlyPheProThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThr 258
DB 781 GTGGTCATGGGGATCCCAACATATGGCACTCCTTCACACTGGCCCTCT---GCAGAAACC 837

QY 259 GlyIleGlyAlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGly 278
DB 838 ACCGGGGGGCCCTCGCTCTGGCCCTGGAGCTGTGGACCCATCACAGAGTCTTCAGGC 897

QY 279 IleTrpAlaTyrTyrGluIleCysThrPheLeuLysAsnGlyAlaThrGlnGlyTrpAsp 298
DB 898 TTCCTGGCCCTATTATGAGATCTGCCAGTCTCTGAAA---GGAGCCCAAGATCACGAGCTC 954

QY 299 AlaProGlnGluValProTyrAlaTyrGlnGlyAsnValTrpValGlyTyrAspAsnIle 318
DB 955 CAGGATCAGAGGTTCCCTTACCGAGTCAAGGGGACCCGGTGGGTGGCTATGATGATGTG 1014
```

RESULT 8

US-09-765-231A-8

; Sequence 8, Application US/09765231A

; Patent No.: US20020119452A1

; GENERAL INFORMATION:

; APPLICANT: Searle/Monsanto

; APPLICANT: Phippard, Deborah

; APPLICANT: Vasanthakamur, Geetha

; APPLICANT: Dotson, Stanton

; APPLICANT: Ma, Xiao-Jun

; TITLE OF INVENTION: Osteoarthritis tissue-derived nucleic acids, polypeptides,

; FILE OF INVENTION: vectors, and cells

; FILE REFERENCE: SO-3221 PR

; CURRENT APPLICATION NUMBER: US/09/765,231A

; CURRENT FILING DATE: 2001-01-18

; NUMBER OF SEQ ID NOS: 82

; SEQ ID NO 8

; LENGTH: 1474

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-765-231A-8

Alignment Scores:

Pred. No.: 4,24e-91 Length: 1474
Score: 873.50 Matches: 181
Percent Similarity: 67.55% Conservative: 73
Best Local Similarity: 48.14% Mismatches: 106
Query Match: 34.86% Indels: 16
DB: 10 Gaps: 5

US-10-004-219B-14 (1-455) x US-09-765-231A-8 (1-1474)

```
QY 1 TyrGlnLeuThrCysTyPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPhe 20
DB 137 TACAAACTGGTTGGTGTACTTTTACCAACTGGGTCCAGGACCGGCAAGAACCCAGGAAAT 196

QY 20 eMetProAsp-AsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPhe-AlaGly 39
DB 197 CACCCCTGAGGATATTCACCCCTTCTATGCTCTCATCTCATCTATTCATTCGCCGAC 256

QY 40 ArgGlnAsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrGlnAlaPhe 59
DB 257 ATCGAAAAACAAGTTTATCATCAAGGACAAGAGTGAAGTGTGCTCTTACCAACCATC 316

QY 60 AsnGly-LeuLysAsnLysAsnSerGlnLeuLysThrLeuLeuAlaIleGlyGlyTrpAs 79
DB 317 AACAGTTCTCAAAACCAAGAATCCCAAACTGAAATTCCTTGTCCATTTGGAGGGTACCT 376

QY 79 nPheGlyThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIle 99
DB 377 GTTTGGTTCCAAAGGGTTCCACCCCTATGTTGGATTCTTCTACATCACGCTTGAATTCAT 436

QY 99 eTherSerValIleLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpG 119
DB 437 TAACTCCATAATCTGTTTCTGAGGAACCAATACTTTGATGAGTGTGATGTAAGCTGGAT 496

QY 119 uTyrProGlySerArgGlySerProGlnAspLysHisLeuPheThrValLeuValG 139
DB 497 CTACCCAGATCAGAAA-----GAAACACTCAT---TTCACCTGTGCTGATTCAT 541
```

```
QY 139 nGluMetArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLysMetVa 159
Db 542 TGAGTTAGCAGAGCCCTTTTCAGAGGACTTCCAAAATCCCAAGAAAGGCTTCTCT 601
QY 159 lThrAla-AlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnL 179
Db 602 GACTCGGGGGGTATCTGCAGGAGGCAAAATGATTGATAACAGCTATCAAGTTGAGAAAC 661
QY 179 eUSeGlnTyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTyrGluG 199
Db 662 TGCAAAAGATCTGGATTTTCATCACTCTGCTTCCTTACTCCATGGTCTTGGGAAA 721
QY 199 ly-----TyrThrGlyGluAsnSerPro-LeuTyrLysTyrProThrAspThrGlySer 216
Db 722 AGCCCTTATCACTGCGCCACACAGCCCTGCTGAGCAAGGGTGGCAGACAGAGGGCCA 781
QY 217 AsnAlaTyrLeuAsnValAspTyrValMetAsnTyrTyrLysAspAsnGlyAlaProAla 236
Db 782 AGCTCTACTACAAATGTGAATATGCTGTGGGTACTGGATACATAAGGGAATGCCATCA 841
QY 237 GluLysLeuIleValGlyPheProThrTyr-GlyHisAsnPheIleLeuSerAsnProSe 256
Db 842 GAGAAGTGGTCAATGGGCATCCCAATATGGGCACTCTTCCACTGGCCCTCT---GC 898
QY 256 xAsnThrGlyIleGlyAlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGl 276
Db 899 AGAAACACCGTGGGGGCCCTCTGCTCTGGCCCTGGAGCTGTGGACCCATCACAGAGTC 958
QY 276 uSerGlyIleTyrAlaTyrTyrGluIleCysThrPheLeuLysAsnGlyAlaThrGlnGl 296
Db 959 TTCAGGCTTCTTGGCCCTATATGAGATCTGCCAGTTCTCTGAAA---GGAGCCAAAGATCAC 1015
QY 296 yTrpAspAlaProGlnGluValProTyrAlaTyrGlnGlyAsnValTyrValGlyTyrAs 316
Db 1016 GCGGCTCCAGGATCACAGGTTCCCTTACGAGTCAAGGGGAACAGTGGTGGGCTATGA 1075
QY 316 pAsnIleLysSerPheAspIleLysAlaGlnTyrLeuLysHisAsnLysPheGlyGlyAl 336
Db 1076 TGATGGAAGAGTATGAGACCAAGGTTTCAGTTCTTTAAAGATTTAAACCTGGAGGAGC 1135
QY 336 aMetValTyrAlaIleAspLeuAspAspPheThrGlyThrPheCysAsnGlnGlyLysPh 356
Db 1136 CATGATCTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1195
QY 356 eProLeuIleSerThrLeuLysLysAlaLeuGly 367
Db 1196 CCTCTTCTCCAGCAGTCAAGAGAGCCTTGGC 1229
```

RESULT 9

```
US-10-218-743-40
; Sequence 40, Application US/10218743
; Publication No. US20030096779A1
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
; AND USES THEREOF
; FILE REFERENCE: AL-2-C3
; CURRENT APPLICATION NUMBER: US/10/218,743
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: US/09/292,225
; PRIOR FILING DATE: 1999-04-15
; PRIOR APPLICATION NUMBER: 60/098,909
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/085,295
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/098,565
; PRIOR FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: 09/062,013
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patent in Ver. 2.0
```

```
; SEQ ID NO 40
; LENGTH: 1470
; TYPE: DNA
; ORGANISM: Dermatophagoides farinae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1470)
US-10-218-743-40

Alignment Scores:
Pred. No.: 2,488-72 Length: 1470
Score: 711.50 Matches: 149
Percent Similarity: 54.17% Conservative: 72
Best Local Similarity: 36.52% Mismatches: 152
Query Match: 28.33% Indels: 35
DB: 9 Gaps: 11

US-10-004-219B-14 (1-455) x US-10-218-743-40 (1-1470)
QY 2 GlnLeuThrCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPheMet 21
Db 43 CGAATCGTATGTTATGTTGGACATGGTCCGTTTATCATAAA---GTTGATCCATACACA 99
QY 22 ProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArgGln 41
Db 100 ATTGAAGATATTGATCCTTTCAAATGTACTCATTTGATGTATGTTTGTCTAAAAATCGAT 159
QY 42 AsnAsnGluIleThr-----IleGluTyr 50
Db 160 GAATCAAAATCACCATTCACAGTTTTTGTATCATTTCAAGATGATGAACCATCACTCATGG 219
QY 51 AsnAspValThrLeuTyrGlnAlaPheAsnGlyLeuLysAsnLysAsnSerGlnLeuLys 70
Db 220 ---GAAACACCGGTATGACGTTTCAACAACTTTGAGATTCGAAGAATCCAGAATTGACC 276
QY 71 ThrLeuLeuAlaIleGlyGlyTyrAsnPheGlyThrAlaPropheThrAlaMetValSer 90
Db 277 ACCATGATTTCATTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 336
QY 91 ThrProGluAsnArgGlnThrPheIleThrSerValIleLysPheLeuArgGlnTyrGlu 110
Db 337 AATCCAAATATCGTCAGCAATTTGTTCAATCAGATTTTGGACTTTTTCGCAAGATACAA 396
QY 111 PheAspGlyLeuAspPheAspTrpGluTyrProGlySerArg---GlySerProProGln 129
Db 397 TTCGATGCGCTAGATTGCGATTGGGAATATCTGTCGATCAGCGTTAGGCAATCTTAAATC 456
QY 130 AspLysHisLeuPheThrValLeuValGlnGluMetArgGluAlaPheGluGlnGluAla 149
Db 457 GATAAACAACAACTATTAACTAGTTAGAGAACTTAAAGAGGCAATTTGAACCTTTTCGGC 516
QY 150 LysGlnIleAsnLysProArgLeuMetValThrAlaAlaValAlaAlaGlyIleSerAsn 169
Db 517 -----TACTTTGTGCTCGCGAGTATCATCCCGGTAAAGATAAA 555
QY 170 IleGlnSerGlyTyrGluIleProGlnLeuSerGlnTyrLeuAspTyrIleHisValMet 189
Db 556 ATTGACGTAGCTTATGAGCTCAAGAAATTTGAACCAATTTGTTGCGATTTGATGATGATG 615
QY 190 ThrTyrAspLeuHisGlySerTrpGluGlyTyrThrGlyGluAsnSerProLeuTyrLys 209
Db 616 ACTTATGATTACCATGGCGGATGGGAAAATGTTTTCGGCCATAATGCTCCGTGTTATATA 675
QY 210 TyrProThrAspThrGlySer---AsnAlaTyrLeuAsnValAspTyrValMetAsnTyr 228
Db 676 CGACCGGATGAACCGATGAATGACACTTACTTCAATGTCACACTACACCATGACTAT 735
QY 229 TrpLysAspAsnGlyAlaProAlaGluLysLeuIleValGlyPheProThrTyrGlyHis 248
Db 736 TATTTGAACAAATGGCGCTACTCGACACAACTTGTATGGGTGTTCCATCTCTATGTCGT 795
QY 249 AsnPheIleLeuSerAsnProSerAsnThrGlyIleGlyAlaProThrSerGlyAlaGly 268
Db 795
```


91 ThrProGluAsnArgGlnThrPheIleThrSerValIleLysPheLeuArgGlnTyrGlu 110
 394 AATCAACACATATCGTCAGCAAAATTTGTTCAATCAGTTTGGACTTTTTCGCAAGAATACAAA 453
 111 PheAspGlyLeuAspPheAspTyrGlySerArg--GlySerProGln 129
 454 TTCGATGCCCTAGATTGGATTGGGAATATCTCGATCAGGTTAGGCAATCTTAAATC 513
 130 AspLysHisLeuPheThrValLeuValGlnMetArgGluAlaPheGluGlnGluAla 149
 514 GATAAACAACAATTAACATTACTAGAGAACCTTAAAGAGCATTTGAACCTTTCCGC 573
 150 LysGlnIleAsnLysProArgLeuMetValThrAlaAlaValAlaAlaGlyIleSerAsn 169
 574 -----TACTTGTTCAGTCGCCGAGTATACCCGGTAAAGATAAA 612
 170 IleGlnSerGlyTyrGluIleProGlnLeuSerGlnTyrLeuAspTyrIleHisValMet 189
 613 ATTAGCAGTAGCTATGAGCTCAAGAATTGAACCAATTGTCGATGGATGAATGATG 672
 190 ThrTyrAspLeuHisGlySerTyrGluGlyTyrThrGlyGluAsnSerProLeuTyrLys 209
 673 ACTTATGATTACCATGGCGATGGGAAATGTTTCGCCCAATATGCTCCGTTGTATAAA 732
 210 TyrProThrAspThrGlySer---AsnAlaTyrLeuAsnValAspTyrValMetAsnTyr 228
 733 CGACCCGATGAAACCGATGAATTGCACACTTACTTCAATGTCAACTACACCATGCACAT 792
 229 TrpLysAspAsnGlyValaProAlaGluLysIleValGlyPheProThrTyrGlyHis 248
 793 TATTTGAACAATGGCGCTACTCGACAAACCTGTGTATGGGTGTTCCATTCTATGGTGT 852
 249 AsnPheIleLeuSerAsnProSerAsnThrGlyIleGlyAlaProThrSerGlyAlaGly 268
 853 GTTTCGAGCATCGAAGATCGAAGCAAAAGTCAAACTGGCGATCCGGCCAAAGCATGTCT 912
 269 ProAlaGlyProTyrAlaLysGlnSerGlyIleTrpAlaTyrTyrGluIleCys---Thr 287
 913 CCTCCTGGTTTTATTACTGGTGAAGAAGGTGTCTCTCATACATCGAATTTGTGTCAGTTA 972
 288 PheLeuLysAsnGlyValaThrGlnGlyTrpAspAlaProGlnGluValProTyrAlaTyr 307
 973 TTTCAGAAAGAAGAATGGCATATTCAATACGATGAATATTACATGTCTCCATACCGGAT 1032
 308 GlnGlyAsnValTyrValGlyTyrAspAsnIleLysSerPheAspIleLysAlaGlnTrp 327
 1033 AATGATAAAATCTGGGTGGTTACGATGATCTGGCTAGTATATCATGCAAGTTGGCCTTT 1092
 328 LeuLysHisAsnLysPheGlyGlyAlaMetValTyrAlaIleAspLeuAspAspPheThr 347
 1093 CTCAAAGAATTCGGCGTCTCTGGCGTTATGATATGCTCATTTGGAAAAACGATGATTCAAA 1152
 348 GlyThrPheCysAsnGlnGlyLysPheProLeuIleSerThrLeuLysLysAlaLeuGly 367
 1153 GGT---CATTCGGGACCG---AAATATCCATTGTTGAACAAAGTTTCACAATATGATCAAT 1206
 368 -----LeuGlnSerAlaSerCysThrAlaProAlaGln 378
 1207 GGTGATGAAGAAGAACTCTTACGATGTCTTTTGGGCCCAAGTACAAACCAACCAACA--- 1263
 379 ProfileGluProIleThrAlaAla 386
 1264 CCAACCCCGCTCAACTACTTCG 1287

RESULT 12
 US-10-218-743-39/c
 ; Sequence 39, Application US/10218743
 ; Publication No. US20030096779A1
 ; GENERAL INFORMATION:
 ; APPLICANT: McCall, Catherine A.
 ; APPLICANT: Hunter, Shirley Wu
 ; APPLICANT: Weber, Eric R.
 ; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS

; TITLE OF INVENTION: AND USES THEREOF
 ; FILE REFERENCE: AL-2-C3
 ; CURRENT APPLICATION NUMBER: US/10/218,743
 ; CURRENT FILING DATE: 2002-08-13
 ; PRIOR APPLICATION NUMBER: US/09/292,225
 ; PRIOR FILING DATE: 1999-04-15
 ; PRIOR APPLICATION NUMBER: 60/098,909
 ; PRIOR FILING DATE: 1998-09-02
 ; PRIOR APPLICATION NUMBER: 60/085,295
 ; PRIOR FILING DATE: 1998-05-13
 ; PRIOR APPLICATION NUMBER: 60/098,565
 ; PRIOR FILING DATE: 1998-04-17
 ; PRIOR APPLICATION NUMBER: 09/062,013
 ; PRIOR FILING DATE: 1998-04-17
 ; NUMBER OF SEQ ID NOS: 49
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 39
 ; LENGTH: 1527
 ; TYPE: DNA
 ; ORGANISM: Dermatophagoides farinae
 US-10-218-743-39

Alignment Scores:
 Pred. No.: 2,638-72 Length: 1527
 Score: 711.50 Matches: 149
 Percent Similarity: 54.17% Conservative: 72
 Best Local Similarity: 36.52% Mismatches: 152
 Query Match: 28.39% Indels: 35
 DB: 9 Gaps: 11

US-10-004-219B-14 (1-455) x US-10-218-743-39 (1-1527)

QY 2 GlnLeuThrCysThrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPheMet 21
 DB 1428 CGAATCGTATGTATGTTGACATGCTCGCTTATCATAAA---GTTGATCCATACACA 1372
 QY 22 ProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArgGln 41
 DB 1371 ATTGAAGATATTGATCCTTTCAAAATGATCTCATTTGATGATGTTTGTCTAAATTCGAT 1312
 QY 42 AsnAsnGluIleThrThr-----ileGluTrp 50
 DB 1311 GAATACAAATACACCATTCAGATTTTGTATCATTCATTCAGATGATAACCATACATCG 1252
 QY 51 AsnAspValThrLeuTyrGlnAlaPheAsnGlyLeuLysAsnLysSerGlnLeuLys 70
 DB 1251 ---GAAACACCGGTATGAGCTTCAACACTTGAGATTGAGATTCAGAAATCCAGATTGACC 1195
 QY 71 ThrLeuLeuAlaIleGlyGlyTrpAsnPheGlyThrAlaProPheThrAlaMetValSer 90
 DB 1194 ACCATGATTTCATTTGGTGGTATGAGGTTTCAGAAAAATATTCGGATATGCGAGCC 1135
 QY 91 ThrProGluAsnArgGlnThrPheIleThrSerValIleLysPheLeuArgGlnTyrGlu 110
 DB 1134 AATCCAAACATATCGTCAGCAATTTGTTCAATCAGTTTTCAGCTTTTTCGAAAGATACAAA 1075
 QY 111 PheAspGlyLeuAspPheAspTrpGluTyrProGlySerArg---GlySerProProGln 129
 DB 1074 TTCATGCTGCTAGATTGATTTGGGATATCTCGATCAGTTAGGCAATCTTAATC 1015
 QY 130 AspLysHisLeuPheThrValLeuValGlnGluMetArgGluAlaPheGluGlnGluAla 149
 DB 1014 GATAAACAACAACTATTAACTATTAGTAGAAGCTTAAAGAGGCAATTTGAACTTCGCGC 955
 QY 150 LysGlnIleAsnLysProArgLeuMetValThrAlaAlaValAlaGlyIleSerAsn 169
 DB 954 -----TACTTGTGACTGCGCAGTATCACCGGTAAAGATAAA 916
 QY 170 IleGlnSerGlyTyrGluIleProGlnLeuSerGlnTyrLeuAspTyrIleHisValMet 189
 DB 915 ATTGACGTAGCTTATGAGCTCAAGAAATTCAGCAATTTGTCATGTTGATGATGTCATG 856
 QY 190 ThrTyrAspLeuHisGlySerTrpGluGlyTyrThrGlyGluAsnSerProLeuTyrLys 209

DB 855 ACTTATGATTACCATGGCGGATGGGAAATGTTTTCGGCCATAATGCTCCGTTGTATAAA 796
 QY 210 TyrProThrAspThrGlySer---AsnAlaTyrLeuAsnValAspTyrValMetAsnTyr 228
 DB 795 CGACCCGATGAACGATGAATTCACACTTACTTCAATGCACTACACCATGACCTAT 736
 QY 229 TrpLysAspAsnGlyAlaProAlaGluLysLeuIleValGlyPheProThrTyrGlyHis 248
 DB 735 TATTTGAACAATGGCGCTACTCGAGACAACTTGTATGGGTGTTCCATTTCTATGGTCT 676
 QY 249 AsnPheIleLeuSerAsnProSerAsnThrGlyIleGlyAlaProThrSerGlyAlaGly 268
 DB 675 GCTTGGAGCATCGAAGATCGAAGCAAACTCAAACTTGGCGATCCGCGCAAGGCAATGCT 616
 QY 269 ProAlaGlyProTyrAlaLysGluSerGlyIleTyrAlaTyrTyrGluIleCys---Thr 287
 DB 615 CTTCTGTTTTTATTACTGCTGGAAGAGTGTCTCTCATACATCGAATTTGTCAGTTA 556
 QY 288 PheLeuLysAsnGlyAlaThrGlnGlyTrpAspAlaProGlnGluValProTyrAlaTyr 307
 DB 555 TTCCAGAAAGAAGAATGCGATATTCATACGATGAATATTACAATGCTCCATACGGATAT 496
 QY 308 GlnGlyAsnValTrpValGlyTyrAspAsnIleLysSerPheAspIleLysAlaGlnTrp 327
 DB 495 AATGATAAAATCTGGGTGTTGATCGATGATCTGGCTAGTATATCATGCAAGTTGGCCTTT 436
 QY 328 LeuLysHisAsnLysPheGlyGlyAlaMetValTrpAlaIleAspLeuAspAspPheThr 347
 DB 435 CTCAAAGAATTGGGCGTCTCTGGGTTATGATATGTCATTGGGCCCCAAGTACCAACCAACA 376
 QY 348 GlyThrPheCysAsnGlnGlyLysPheProLeuIleSerThrLeuLysLysAlaLeuGly 367
 DB 375 GGT---CATTTGGGACCG---AAATATCCATTTGTTGAACAAAGTTTCAATATATGCAAT 322
 QY 368 -----LeuGlnSerAlaSerCysThrAlaProAlaGln 378
 DB 321 GGTGATGAAAGAAGAACTCTTAGCAATGCTTTTGGGCCCAAGTACCAACCAACA--- 265
 QY 379 ProIleGluProIleThrAlaAla 386
 DB 264 CAACCCCGCTCAACTACTTCG 241
 RESULT 13
 US-10-218-743-34
 ; Sequence 34, Application US/10218743
 ; Publication No. US20030096779A1
 ; GENERAL INFORMATION:
 ; APPLICANT: McCall, Catherine A.
 ; APPLICANT: Hunter, Shirley Wu
 ; APPLICANT: Weber, Eric R.
 ; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
 ; TITLE OF INVENTION: AND USES THEREOF
 ; FILE REFERENCE: AL-2-C3
 ; CURRENT APPLICATION NUMBER: US/10/218,743
 ; CURRENT FILING DATE: 2002-08-13
 ; PRIOR APPLICATION NUMBER: US/09/292,225
 ; PRIOR FILING DATE: 1999-04-15
 ; PRIOR APPLICATION NUMBER: 60/098,909
 ; PRIOR FILING DATE: 1998-09-02
 ; PRIOR APPLICATION NUMBER: 60/085,295
 ; PRIOR FILING DATE: 1998-05-13
 ; PRIOR APPLICATION NUMBER: 60/098,565
 ; PRIOR FILING DATE: 1998-04-17
 ; PRIOR APPLICATION NUMBER: 09/062,013
 ; PRIOR FILING DATE: 1998-04-17
 ; NUMBER OF SEQ ID NOS: 49
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 34
 ; LENGTH: 1621
 ; TYPE: DNA
 ; ORGANISM: Dermatophagoides farinae
 ; FEATURE:

```
NAME/KEY: CDS
LOCATION: (14)..(1540)
US-10-218-743-34

Alignment Scores:
Pred. No.: 2,89e-72 Length: 1621
Score: 711.50 Matches: 149
Percent Similarity: 54.17% Conservative: 72
Best Local Similarity: 36.52% Mismatches: 152
Query Match: 28.39% Indels: 35
DB: Gaps: 11

US-10-004-219B-14 (1-455) x US-10-218-743-34 (1-1621)
QY 2 GlnLeuThrCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPheMet 21
DB 113 CGAATCGTATGTTATGCGAATCGTCCGTTTATCATATAA---GTTGATCCATACACA 169
QY 22 ProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArgGln 41
DB 170 ATTGAAGATATGATCTCTTCAAAATGATCTCATTTGATGTTGTTTGGTAAATCGAT 229
QY 42 AsnAsnGluIleThrThr-----lleGluTrp 50
DB 230 GAATACAAATACACCATTCATCAAGTTTGTGATCCATTTCAAGATGATTAACCATCACTCGG 289
QY 51 AsnAspValThrLeuTyrGlnAlaPheAsnGlyLeuLysAsnLysSerGlnLeuLys 70
DB 290 ---GAAACACCGGATGAACCTTTCAACAACTTGAGATTGAGATTGAGATTGAGATTGAGC 346
QY 71 ThrLeuLeuAlaIleGlyGlyTyrAsnPheGlyThrAlaProPheThrAlaMetValSer 90
DB 347 ACCATGATTTCAATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 406
QY 91 ThrProGluAsnArgGlnThrPheIleThrSerValIleLysPheLeuArgGlnTyrGlu 110
DB 407 AATCCACATATCGTCAGCAATTTGTTCAATCAGTTTGGACTTTTGGCAAGATAAACA 466
QY 111 PheAspGlyLeuAspPheAspTrpGluTyrProGlySerArg---GlySerProGln 129
DB 467 TTCGATGGCTAGATTGGATTGGGAATATCCCTGGATCAGGTTAGGCAATCTCAAAATC 526
QY 130 AspLysHisLeuPheThrValLeuValGlnGluMetArgGluAlaPheGluGlnGluAla 149
DB 527 GATAAACAACACTATTAACTAGTTAGAGACTTAAAGAGCAATTTGAACCTTTCCGCG 586
QY 150 LysGlnIleAsnLysProArgLeuMetValThrAlaAlaValAlaAlaGlyIleSerAsn 169
DB 587 -----TACTTTGTTGACTCGCGCAGTATCACCCCGTAAAGATAA 625
QY 170 IleGlnSerGlyTyrGluIleProGlnLeuSerGlnTyrLeuAspTyrIleHisValMet 189
DB 626 ATTGAGTGTATGAGCTCAAGAAATGAACCAATTTGCGATTGGGATGGAATGTCATG 685
QY 190 ThrTyrAspLeuHisGlySerTrpGluGlyTyrThrGlyGluAsnSerProLeuTyrLys 209
DB 686 ACTTATGATTACATCGCGGATGGGAATGTTTCGGCCATATGCTCGTTGTATATA 745
QY 210 TyrProThrAspThrGlySer---AsnAlaTyrLeuAsnValAspTyrValMetAsnTyr 228
DB 746 CGACCCGATGAACCGGATGAATTCACACTTACTTCAATGTCAACTACACCATGCACTAT 805
QY 229 TrpLysAspAsnGlyAlaProAlaGluLysLeuIleValGlyPheProThrTyrGlyHis 248
DB 806 TATTTGAACAAATGGCGCTACTCGAGACAACTTTGATGGGTGTTCCATTTCAATGTCGT 865
QY 249 AsnPheIleLeuSerProSerAsnThrGlyIleGlyAlaProThrSerGlyAlaGly 268
DB 866 GCTTGGGATCGAAGATCGAAGCAAACTTGGCGATCCGCCAAAGCATGTCT 925
QY 269 ProAlaGlyProTyrAlaLysGluSerGlyIleTrpAlaTyrTyrGluIleCys---Thr 287
DB 926 CCTCCTGGTGTATTACTGGTGAAGAGGTGTTCTCTCATACATCGAATTTGTGTCAGTTA 985
```

RESULT 14

```
US-10-218-743-36/c
Sequence 36, Application US/10218743
Publication No. US20030096779A1
GENERAL INFORMATION:
APPLICANT: McCall, Catherine A.
APPLICANT: Hunter, Shirley Wu
APPLICANT: Weber, Eric R.
TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
FILE REFERENCE: AL-2-C3
CURRENT FILING DATE: 2002-08-13
PRIOR APPLICATION NUMBER: US/10/218,743
CURRENT FILING DATE: 2002-08-13
PRIOR APPLICATION NUMBER: US/09/292,225
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/098,909
PRIOR FILING DATE: 1998-09-02
PRIOR APPLICATION NUMBER: 60/085,295
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/098,565
PRIOR FILING DATE: 1998-04-17
PRIOR APPLICATION NUMBER: 09/062,013
PRIOR FILING DATE: 1998-04-17
NUMBER OF SEQ ID NOS: 49
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 36
LENGTH: 1621
TYPE: DNA
ORGANISM: Dermatophagoides farinae
US-10-218-743-36

Alignment Scores:
Pred. No.: 2,89e-72 Length: 1621
Score: 711.50 Matches: 149
Percent Similarity: 54.17% Conservative: 72
Best Local Similarity: 36.52% Mismatches: 152
Query Match: 28.39% Indels: 35
DB: Gaps: 11

US-10-004-219B-14 (1-455) x US-10-218-743-36 (1-1621)
QY 2 GlnLeuThrCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPheMet 21
DB 1509 CGAATCGTATGTTATGCGAATCGTCCGTTTATCATATAA---GTTGATCCATACACA 1453
QY 22 ProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArgGln 41
DB 1452 ATTGAAGATATGATCTCTTCAAAATGATGTTTCAATTTGATGTTTGGTAAATCGAT 1393
```

```
QY 42 AsnAsnGluIleThrThr-----IleGluTrp 50
Db 1392 GAATACAAATACACCACTCAAGTTTGTATCCATTTCAAGATGATAACCAATCACTG 1333
QY 51 AsnAspValThrLeuTyrglnAlaPheAsnGlyLeuLysAsnLysAsnSerGlnLeuLys 70
Db 1332 ---GAAACACCGGTATGAAGCTTTCAACACTTTGAGATTGAAGATCCAGAAATGACC 1276
QY 71 ThrLeuLeuAlaIleGlyTrpAsnPheGlyThrAlaProPheThrAlaMetValSer 90
Db 1275 ACCATGATTTCATCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1216
QY 91 ThrProGluAsnArgGlnThrPheIleThrSerValIleLysPheLeuArgGlnTrpGlu 110
Db 1215 AATCCAACATATCGTCAGCAATTTGTTCAATCAGTTTGGAGCTTTTGGCAAGATACAAA 1156
QY 111 PheAspGlyLeuAspPheAspTrpGluTyrglnTyrglnTyrglnTyrglnTyrgln 129
Db 1155 TTTGATGGCTAGATTGGAATTTGGAATTTGGAATTTGGAATTTGGAATTTGGAATTT 1096
QY 130 AspLysHisLeuPheThrValLeuValGlnGluMetArgGluAlaPheGluGlnGluAla 149
Db 1095 GATAACAAACTATTTAACATTAGTTAGAGAACTTTAAGAGGCAATTTGAACCTTTCCGC 1036
QY 150 LysGlnIleAsnLysProArgLeuMetValThrAlaAlaValAlaAlaGlyIleSerAsn 169
Db 1035 -----TACTTTGTTGCTGCGCAGTATCACCGGTAAAGATAAA 997
QY 170 IleGlnSerGlyTyrglnLeuProGlnLeuSerGlnTyrglnLeuAspTyrglnHisValMet 189
Db 996 ATTGACGTAGCTTATGAGCTTAAAGAACTTTGAACCAATTTGCGATTGGATGAATGTCATG 937
QY 190 ThrTyrglnLeuHisGlySerTrpGluGlyTyrglnTyrglnTyrglnTyrglnTyrgln 209
Db 936 ACTTATGATTACATGCGGATGGGAATGTTTTCGGCCATAATGCTCCGCTGGTGTATATA 877
QY 210 TyrProThrAspThrGlySer---AsnAlaTyrglnLeuAsnValAspTyrglnValMetAsnTyrgln 228
Db 876 CGACCCGATGAACCGATGAATGTCACACTTACTTCAATGTCACTACACCATGACACTAT 817
QY 229 TrpLysAspAsnGlyAlaProAlaGlnLysLeuIleValGlyPheProThrTyrglnHis 248
Db 816 TATTTGAACAATGGGCTACTCGACAACTTGTATGGGTGTTCCATTTCTATGTCGT 757
QY 249 AsnPheIleLeuSerAsnProSerAsnThrGlyIleGlyAlaProThrSerGlyAlaGly 268
Db 756 GCTTGGAGCATCGAAGATCGAAGCAAGTCAAACTTGGCGATCCGCGCCAAAGGCATGCT 697
QY 269 ProAlaGlyProTyrglnAlaLysGluSerGlyIleTrpAlaTyrglnTyrglnTyrgln 287
Db 696 CCTCTGTTTATTTACTGTTGAGAGGTTCTCTCATACATCGAATTTGTCAGTTA 637
QY 288 PheLeuLysAsnGlyAlaThrGlnGlyTrpAspAlaProGlnGluValProTyrglnAlaTyrgln 307
Db 636 TTTCCGAAAGAAGATGGCATATTCATACGATGAATATTAATGCTCTCCATACGATAT 577
QY 308 GlnGlyAsnValTrpValGlyTyrglnPheAsnIleLysSerPheAspIleLysAlaGlnTrp 327
Db 576 AATGATAAATCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 517
QY 328 LeuLysHisAsnLysPheGlyGlyAlaMetValTrpAlaIleAspLeuAspPheThr 347
Db 516 CTCAGAAGATTGGCGCTCTCTCGCTTATGATATGTTGTCATTGGAAACGATGATTTCAA 457
QY 348 GlyThrPheCysAsnGlnGlyLysPheProLeuIleSerThrLeuLysLysAlaLeuGly 367
Db 456 GGT---CATTTGGCGCG---AAATATCATTTGTTGAACAAAGTTCACAAATATGATCAAT 403
QY 368 -----LeuGlnSerAlaSerCysThrAlaProAlaGln 378
Db 402 GGTGATGAAGAAGACTTTTACGAATGCTTTTGGGCCCAAGTACAAACCAACCAACA--- 346
```

```
QY 379 ProIleGluProIleThrAlaAla 386
Db 345 CCAACCAACCGCTCACTACTCTG 322
RESULT 15
US-10-218-743-20
; Sequence 20, Application US/10218743
; Publication No. US20030096779A1
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
; FILE REFERENCE: AL-2-C3
; CURRENT APPLICATION NUMBER: US/10/218,743
; PRIOR APPLICATION NUMBER: US/09/292,225
; PRIOR FILING DATE: 1999-04-15
; PRIOR APPLICATION NUMBER: 60/098,909
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/085,295
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/098,565
; PRIOR FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: 09/062,013
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 1608
; TYPE: DNA
; ORGANISM: Dermatophagoides farinae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1608)
US-10-218-743-20
Alignment Scores:
Pred. No.: 6,35e-72 Length: 1608
Score: 708.50 Matches: 149
Percent Similarity: 53.98% Conservative: 75
Best Local Similarity: 35.90% Mismatches: 154
Query Match: 28.27% Indels: 37
Dbs: 9 Gaps: 10
US-10-004-219B-14 (1-455) x US-10-218-743-20 (1-1608)
QY 2 GlnLeuThrCysTyrglnPheThrAsnTrpAlaGlnTyrglnTyrglnLeuGlyArgPheMet 21
Db 43 AGAATTGTTTGTATGTTGGAACATGTCCTGATATATCAATAA---GTTGATCCATACACT 99
QY 22 ProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrglnAlaPheAlaGlyArgGln 41
Db 100 ATCGAAGATATTGATCCATTCAGTGTACACATTAATGATGTTTCGCTAAATTCAT 159
QY 42 AsnAsnGluIleThrThr-----IleGluTrp 50
Db 160 GAATACAAATACACATTCAGTTTTCATCTTACCAAGATGATAACCAATCACTG 219
QY 51 AsnAspValThrLeuTyrglnAlaPheAsnGlyLeuLysAsnLysAsnSerGlnLeuLys 70
Db 220 ---GAAACACCGGTATGAAGCTTTCAACACTTTCGATTCGATTCGATTCGATTCGATTC 276
QY 71 ThrLeuLeuAlaIleGlyTrpAsnPheGlyThrAlaProPheThrAlaMetValSer 90
Db 277 ACCATGATTTCATCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 336
QY 91 ThrProGluAsnArgGlnThrPheIleThrSerValIleLysPheLeuArgGlnTrpGlu 110
Db 337 AATCCAACATATCGTCAGCAATTTGGAATTTGGAATTTGGAATTTGGAATTTGGAATTT 396
QY 111 PheAspGlyLeuAspPheAspTrpGluTyrglnTyrglnTyrglnTyrglnTyrgln 129
```

Db 397 TTCCGCGGTCTAGATTGGATTGGAGTATCTCGATTCGATTGGGTAAACCGAAATC 456
QY 130 AspLysHisLeuPheThrValLeuGlnGluMetArgGluAlaPheGluGlnGluAla 149
Db 457 GATAAACAACAACTATTGGCTTTGGTTAGAGAACTTAAAGACGCTTTGAACCTCATGGC 516
QY 150 LysGlnIleAsnLysProArgLeuMetValThrAlaAlaValAlaAlaGlyIleSerAsn 169
Db 517 -----TACTTGTTCACCTGCTGCAGTATCACCAGGTAAGACAAA 555
QY 170 IleGlnSefGlyTyrgluLeProGlnLeuSerGlnTyrlLeuAspTyrlleHisValMet 189
Db 556 ATCCGACCGAGCTTATGATATCAAGAATGAACAAATTTTCGATTGGATGAATGTCTATG 615
QY 190 ThrTyrlAspLeuHisGlySerTrpGluGlyTyrlThrGlyGluAsnSerProLeuTyrlLys 209
Db 616 ACATATGATTACCCAGGTGGATGGGAAACTTTTACGGTCACAAATGCTCCGTGTATATAA 675
QY 210 TyrProThrAspThrGlySer---AsnAlaTyrlLeuAsnValAspTyrlValMetAsnTyrl 228
Db 676 CGACCCAGATGAACCTGATGAGTTGCACACTTACTTCAATGTCACTACACCATGCACATAT 735
QY 229 TrpLysAspAsnGlyAlaProAlaGluLysLeuIleValGlyPheProThrTyrlGlyHis 248
Db 736 TATTGAACAATGGTCCACGAGACAAATGGTAATGGGTGTTCTCATTTCTATGGCCGT 795
QY 249 AsnPheIleLeuSerAsnProSerAsnThrGlyIleGlyAlaProThrSerGlyAlaGly 268
Db 796 GCTTGGAGCAATGAAGATCGAAGCAAACTCAAACTTGGAGATCCAGCCAAAGGCATGTCG 855
QY 269 ProAlaGlyProTyrlAlaLysGluSerGlyIleTrpAlaTyrlTyrlGluIleCys---Thr 287
Db 856 CCCCAGGTTTCAATTTCTGGTGAAGAGGTGCTCTCATATATAGAAATGTGTCATTTG 915
QY 288 PheLeuLysAsnGlyAlaThrGlnGlyTyrlAspAlaProGlnGluValProTyrlAlaTyrl 307
Db 916 TTTCAAAAGAGAAGATGGCATATCCAATACGATGAATATTACAATGCTCCATATGGTTAC 975
QY 308 GlnGlyAsnValTrpValGlyTyrlAspAsnIleLysSerPheAspIleLysAlaGlnTrp 327
Db 976 AATGATAAAATCTGGGTGCGTTACGATGATCTGCGCAGTATATCATGCAAGTTGGCTTTC 1035
QY 328 LeuLysHisAsnLysPheGlyGlyAlaMetValTrpAlaIleAspLeuAspPheThr 347
Db 1036 CTGAAGAATTAGCGCTTCTGCTGTCATGTTTGTTCATTCGAAATGATGATTTCAA 1095
QY 348 GlyThrPheCysAsnGlnGlyLysPheProLeuIle----- 359
Db 1096 GGT---CACTGCGGACCG---AAAAATCCATTGTTGAACAAAGTTCATATATGATTAAT 1149
QY 360 -----SerThrLeuLysLysAlaLeuGlyLeuGlnSerAlaSerCysThrAla 375
Db 1150 GCGGATGAAGAAGAACTTTTCGAATGCAATTTGGTCCAGTACACGACCACTCCCA 1209
QY 376 ProAlaGlnProIleGluProIleThrAlaAlaProSerGlySer 390
Db 1210 ACGACGACCCACACCCCGACTACACGCCCAACTCTTCT 1254

Search completed: June 30, 2003, 03:58:41
Job time : 190.913 secs

1	1011.5	40.4	721	1	OGP MOUSE.	Q62010	mus musculus
2	976.5	39.0	383	1	C3L1_HUMAN	P36222	homo sapien
3	970.5	38.7	390	1	C3L2_HUMAN	Q15782	homo sapien
4	961.5	38.4	671	1	OGP MESAU	Q05957	mesocricetu
5	959.5	38.3	678	1	OGP HUMAN	Q12889	homo sapien
6	954.5	38.1	623	1	OGP PAPAN	P36718	papio anubi
7	943.5	37.6	537	1	OGP BOVIN	Q28042	bos taurus
8	941.5	37.6	539	1	OGP SHEEP	Q28542	ovis aries
9	936.5	37.4	527	1	OGP PIG	Q28990	sus scrofa
10	885	35.3	617	1	CHIT CAEBL	Q11174	caenorhabdi
11	879.5	35.1	381	1	C3L1_MOUSE	Q61362	mus musculus
12	863	34.4	554	1	CHIT MANSE	P36362	manduca sex
13	786.5	31.4	504	1	CHIT BRUMA	P29030	brugia mala
14	465.5	18.6	423	1	CH1L APHAL	P23470	aphanoclad
15	463.5	18.5	699	1	CH1L_BACCI	P20533	bacillus ci
16	434.5	17.3	423	1	CH1L TRIHA	P48827	trichoderma
17	407	16.2	427	1	CH1L COCIM	P54196	coccidioid
18	344.5	13.7	499	1	CHIB SERNA	P11797	serratia ma
19	343	13.7	1046	1	CHIB VIFBU	P66156	vibrio furn
20	342.5	13.7	563	1	CHIA_SERMA	P07254	serratia ma
21	342	13.6	820	1	CHIA_ALTSO	P2823	alteromonas
22	323.5	12.9	619	1	CHIT STRLI	P36909	streptomyce
23	319	12.7	550	1	CHIT NPVOF	Q10363	orgyia pseu
24	312.5	12.5	610	1	CHIT STRPL	P11220	streptomyce
25	303	12.1	551	1	CHIT NPVAC	P41684	autographa
26	236.5	9.4	1146	1	KTXA_KLULA	P09805	kluveryomyc
27	198.5	7.9	367	1	DIAC_RAT	Q01460	rattus norv
28	196.5	7.8	385	1	DIAC_HUMAN	Q01459	homo sapien
29	154.5	6.2	427	1	YAAH_BACSU	P37531	bacillus su
30	142	4.9	524	1	CHIT_BACCI	P27050	bacillus ci
31	121.5	4.8	597	1	CHIX_STROI	Q05638	streptomyce
32	113	4.5	36	1	C3L1_BOVIN	P30922	bos taurus
33	111	4.4	491	1	YKCS CAEBL	P41996	caenorhabdi

Query Match 38.3%; Score 959.5; DB 1; Length 678;
Best Local Similarity 49.3%; Pred. No. 1.8e-62;
Matches 187; Conservative 56; Mismatches 129; Indels 7; Gaps 3;

QY 1 YLTCTYTNWQYRPGCLGRFMPDNIDPCLCTHLIYAFAGQNNETITIEWND-VTLYQAF 59
DB 22 HKLVCFYTNWHSRPGPASIPLDLPFLCTHLIYAFASMMNNQIVAKDQEKILYPEF 81

QY 60 NGLKNKNSQLKTLAIGGNFCTAPMTAVSPENROTFTITSVIFKLOYEFGDGLDFWE 119
DB 82 NKLKERNRELKTLISIGGNFCTSRFTMTLSTFANREKFIASVISLRLTHDPLDPL 141

QY 120 YPGSRGSPQDKHLFTVLVQEMREAFEOBAKQINPRLMVAAGISNIQSGYEIPOL 179
DB 142 YPGLRGSPMHDRTWFLFIEELLFAPRKEALLTMRPRLLSAAVSGVPHIVQTSYDVRL 201

QY 180 SQYLDYIHWYVYDLHSGWEGYGENSPLYKYPTDGTGSNAYLVNDVYMWYKNGGAPAEKL 239
DB 202 GRLLDFINVLSDYDLHSGWERTGHNPLSPFSDPKSSA-----YAMNYWRKLGAPSEKL 256

QY 240 IVGFPTYGHNFTILSNPNTGICAPTSGAGPAGPYAKESGIWAYYEICTFLKNGATQGWDA 299
DB 257 IMGIFTYGTFRLLKASKNGLOARATGASPCKYTKQBFLEYFICSFVW-GAKKHWD 315

QY 300 PQEVPYAYQGNVWGYDNIKSFIDIKAOMLKHNKFGGAMWAIDLDDFTGTCNQKGFPLI 359
DB 316 YQYVPYANKGEWGVYDIAISFSYKAWFIRREHFGGAMVWTLDMDDVRCGTGTPFPLV 375

QY 360 STLKALGLQASCTAPAQ 378
DB 376 YVINDILVRAEFSSLSLPQ 394

RESULT 6
OGP_PAPAN STANDARD; PRT; 623 AA.
AC P36718;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Oviduct-specific glycoprotein precursor (Oviductal glycoprotein)
DE (Oviductin) (Estrogen-dependent oviduct protein).
GN OVGPI OR OGP.
OS Papio anubis (Olive baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Papio.
OX NCBI_TaxID=9555;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Oviduct;
RX MEDLINE=98244335; PubMed=9584944;
RA Verhage H.G., Fazleabas A.T., Mavrogianis P.A., O'Day-Bowman M.B.,
Donnelly K.M., Arias E.B., Jaffe R.C.;
RT "The baboon oviduct: Characteristics of an oestradiol-dependent
oviduct-specific glycoprotein."
RL Hum. Reprod. Update 3:541-552(1997).
RN [2]
RP PRELIMINARY PARTIAL SEQUENCE FROM N.A.
RC TISSUE=Oviduct;
RX MEDLINE=91367180; PubMed=1716345;
RA Donnelly K.M., Fazleabas A.T., Verhage H.G., Mavrogianis P.A.,
Jaffe R.C.;
RT "Cloning of a recombinant complementary DNA to a baboon (Papio
anubis) estradiol-dependent oviduct-specific glycoprotein."
RL Mol. Endocrinol. 5:356-364(1991).
CC -!- FUNCTION: BINDS TO OCYOTE ZONA PELLUCIDA IN VIVO. MAY PLAY A ROLE
IN THE FERTILIZATION PROCESS AND/OR EARLY EMBRYONIC DEVELOPMENT.
CC -!- SUBCELLULAR LOCATION: Secretory granules.
CC -!- TISSUE SPECIFICITY: OVIDUCT.
CC -!- DEVELOPMENTAL STAGE: AT THE TIME OF OVULATION.
CC -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M59903; AAB39765.1; -;
DR PIR; A37954; A37954.
DR InterPro; IPR001579; Chitinase_18/2.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
KW Glycoprotein; Fertilization; Signal.
FT SIGNAL 1 21 BY SIMILARITY.
FT CHAIN 22 623 OVIDUCT-SPECIFIC GLYCOPROTEIN.
FT CARBOHYD 402 402 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 441 441 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 623 AA; 69291 MW; 9E21CE481FFF1268 CRC64;
Query Match 38.1%; Score 954.5; DB 1; Length 623;
Best Local Similarity 48.8%; Pred. No. 3.6e-62;
Matches 185; Conservative 59; Mismatches 128; Indels 7; Gaps 3;

QY 1 YLTCTYTNWQYRPGCLGRFMPDNIDPCLCTHLIYAFAGQNNETITIEWND-VTLYQAF 59
DB 22 HKLVCFYTNWHSRPGPASIPLDLPFLCTHLIYAFASMMNNQIVAKDQEKILYPEF 81

QY 60 NGLKNKNSQLKTLAIGGNFCTAPMTAVSPENROTFTITSVIFKLOYEFGDGLDFWE 119
DB 82 NKLKERNRELKTLISIGGNFCTSRFTMTLSTFANREKFIASVISLRLTHDPLDPL 141

QY 120 YPGSRGSPQDKHLFTVLVQEMREAFEOBAKQINPRLMVAAGISNIQSGYEIPOL 179
DB 142 YPGLRGSPMHDRTWFLFIEELLFAPRKEALLTMRPRLLSAAVSGVPHIVQTSYDVRL 201

QY 180 SQYLDYIHWYVYDLHSGWEGYGENSPLYKYPTDGTGSNAYLVNDVYMWYKNGGAPAEKL 239
DB 202 GRLLDFINVLSDYDLHSGWERTGHNPLSPFSDPKSSA-----YAMNYWRKLGAPSEKL 256

QY 240 IVGFPTYGHNFTILSNPNTGICAPTSGAGPAGPYAKESGIWAYYEICTFLKNGATQGWDA 299
DB 257 IMGIFTYGTFRLLKASKNGLOARATGASPCKYTKQBFLEYFICSFVW-GAKKHWD 315

QY 300 PQEVPYAYQGNVWGYDNIKSFIDIKAOMLKHNKFGGAMWAIDLDDFTGTCNQKGFPLI 359
DB 316 YQYVPYANKGEWGVYDIAISFSYKAWFIRREHFGGAMVWTLDMDDVRCGTGTPFPLV 375

QY 360 STLKALGLQASCTAPAQ 378
DB 376 YVINDILVRAEFSSLSLPQ 394

RESULT 7
OGP_BOVIN STANDARD; PRT; 537 AA.
AC Q28042;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Oviduct-specific glycoprotein precursor (Oviductal glycoprotein)
DE (Oviductin) (Estrogen-dependent oviduct protein) (fragment).
GN OVGPI OR OGP.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 19-47.

RC TISSUE=Oviduct;
RX MEDLINE=94257768; PubMed=8199272;
RA Sendai Y., Abe H., Kikuchi M., Satch T., Hoshi H.;
RT "Purification and molecular cloning of bovine oviduct-specific
glycoprotein";
RL Biol. Reprod. 50:927-934 (1994).
CC -!- FUNCTION: BINDS TO OOCYTE ZONA PELLUCIDA IN VIVO. MAY PLAY A ROLE
CC IN THE FERTILIZATION PROCESS AND/OR EARLY EMBRYONIC DEVELOPMENT.
CC -!- SUBCELLULAR LOCATION: Secretory granules.
CC -!- TISSUE SPECIFICITY: OVIDUCT.
CC -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D16639; BAA04065.1; -
DR InterPro; IPR001579; Chitinase 18/2.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR PROSITE; PS01095; CHITINASE 18; FALSE_NEG.
KW Glycoprotein; Fertilization; Signal.
FT NON_TER 1
FT SIGNAL <1 18
FT CHAIN 19 537 OVIDUCT-SPECIFIC GLYCOPROTEIN.
FT CARBOHYD 399 399 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 537 AA; 59617 MW; CFCDEE6F0212D791 CRC64;

Query Match 37.6%; Score 943.5; DB 1; Length 537;
Best Local Similarity 49.8%; Pred. No. 1.9e-61;
Matches 182; Conservative 53; Mismatches 125; Indels 7; Gaps 3;

QY 1 YOLTCTFTNWAQYRPGLRGFMNDPCLCTHLIYAFAGRQNNETITIEWND-VTLYQAF 59
DB 19 HKLVCTFTNWAQYRPGASLPDLDFLCTHLVAFASNNQIVPKDPQDEKILYPEF 78

QY 60 NGLKNNQSLKTLATAGNFGTAPMTAMVSTENRQTITSVIKELRQVDFDGLDFWE 119
DB 79 NKLKERNGLKTLISGWNFGTAVRTTMTLSTFNRERFVSSVIALRTHTGFDGLDFPL 138

QY 120 YGSRGSPQDKHLFTVLQEMREAEQAKQINKPLMTAAVAAGISNIQSGYEIPOL 179
DB 139 YPLRGSPARDRTTFVLEELQAFKNEAQLTRPRLLLSAAVSGDPHVVKAYEARLL 198

QY 180 SQYLDIHYMTYDLHGSWEGYTGENSEPLYKYPTDGTGSNAYLVNDYVMYWKONGAPAEKL 239
DB 199 GRLLDFISLVSLDHLGSWEKVTGHNPSLPFGDPKSSA-----YANNYRQLGVPEKL 253

QY 240 IVGFPTTHNFILNSPNTGIGAPTSAGAGPYAKESGIWAYEICTFLKNGATGWDA 299
DB 254 LMGLPTYGRTHLLKASQNELRAQAVGSPGKYTKQAGFLAYEITCCFVRR-AKRRWIN 312

QY 300 PQEVPYAGNVWGVNDIKSPDIKAQWLKHNKFGAMVWADLDPTGTCNQGKFPPLI 359
DB 313 DQVVPYAFKGEWGVYDDAISFGYKAFFIKREHFGGAMVWTLDDDFRGYFCGTGFPPLV 372

QY 360 STLKAL 366
DB 373 HTLNNLL 179

RESULT 8
OGP SHEEP
ID_OGP SHEEP STANDARD; PRT; 539 AA.
AC Q28542; Q28543;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE Oviduct-specific glycoprotein precursor (Oviductal glycoprotein)
DE (Oviductin) (Estrogen-dependent oviduct protein) (Estrus-associated
DE oviductal glycoprotein) (OEGP).
GN OVGPI OR OGP.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Artiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 22-39.
RX TISSUE=Oviduct;
RX MEDLINE=95269691; PubMed=7750470;
RA Desouza M.M., Murray M.K.;
RT "An estrogen-dependent secretory protein, which shares identity with
RT chitinase, is expressed in a temporally and regionally specific
RT manner in the sheep oviduct at the time of fertilization and embryo
RT development.";
RT Endocrinology 136:2485-2496 (1995).
RN [2]
RP SEQUENCE OF 10-539 FROM N.A.
RC STRAIN=Merino; TISSUE=Oviduct;
RX MEDLINE=96329120; PubMed=8726871;
RA Marshall J.T.A., Nancarrow C.D., Brownlee A.G.;
RT "Cloning and sequencing of a cDNA encoding an ovine
RT estrus-associated oviductal protein";
RL Reprod. Fertil. Dev. 8:305-310 (1996).
CC -!- FUNCTION: BINDS TO OOCYTE ZONA PELLUCIDA IN VIVO. MAY PLAY A ROLE
CC IN THE FERTILIZATION PROCESS AND/OR EARLY EMBRYONIC DEVELOPMENT.
CC -!- SUBCELLULAR LOCATION: Secretory granules.
CC -!- TISSUE SPECIFICITY: OVIDUCT.
CC -!- DEVELOPMENTAL STAGE: LEVELS ARE HIGHEST IN THE FIMBRIA AND AMPULLA
CC AT ESTRUS AND ON DAY 1 OF PREGNANCY, WHEN GAMETE TRANSPORT AND
CC FERTILIZATION OCCURS IN THE E2-DOMINATED FALLOPIAN TUBE. LEVELS
CC DECLINE SIGNIFICANTLY ON DAY 2 AND UNDERGO A FURTHER SIGNIFICANT
CC REDUCTION ON DAY 3 OF PREGNANCY COINCIDENT WITH TRANSPORT OF THE
CC EMBRYO FROM THE OVIDUCT TO THE UTERUS, A REPRODUCTIVE STAGE
CC ASSOCIATED WITH RISING PROGESTERONE LEVELS.
CC -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U16719; AAC48471.1; -
DR EMBL; U17988; AAB01052.1; -
DR InterPro; IPR001579; Chitinase 18/2.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR PROSITE; PS01095; CHITINASE 18; FALSE_NEG.
KW Glycoprotein; Fertilization; Signal.
FT SIGNAL 1 21
FT CHAIN 22 539 OVIDUCT-SPECIFIC GLYCOPROTEIN.
FT CARBOHYD 402 402 N-LINKED (GLCNAC...) (POTENTIAL).
FT CONFLICT 13 13 M -> V (IN REF. 2).
FT CONFLICT 122 122 K -> N (IN REF. 2).
FT CONFLICT 282 282 A -> V (IN REF. 2).
FT CONFLICT 375 375 A -> V (IN REF. 2).
FT CONFLICT 484 484 R -> H (IN REF. 2).
FT CONFLICT 520 520 I -> T (IN REF. 2).
SQ SEQUENCE 539 AA; 59535 MW; F35000269987C193 CRC64;

Query Match 37.6%; Score 941.5; DB 1; Length 539;
Best Local Similarity 48.8%; Pred. No. 2.7e-61;
Matches 179; Conservative 59; Mismatches 122; Indels 7; Gaps 3;

QY 1 YOLTCTFTNWAQYRPGLRGFMNDPCLCTHLIYAFAGRQNNETITIEWND-VTLYQAF 59
DB 19 HKLVCTFTNWAQYRPGASLPDLDFLCTHLVAFASNNQIVPKDPQDEKILYPEF 78

Db 22 HKLVCFYTNWAFSRPGSASILPRDLDPFLCTHLVFAFASMMNNQIVPKDPLKILYPEF 81
QY 60 NGLKNKNSQLTKLLAIGWNGFTAPFTAMVSTPENROTEITSVIRKLOYEDGLDFDWE 119
Db 82 NKLKERNRGLTKLLSVGGWNGFTSRFTKMLSTFSNRERVKSVIALLRHFGDGLDLPL 141
QY 120 YPGSRGSPQDKHLFTVLVOEMREAFEOBAKQINPKRLMVTAAVAGISNIOGSVEIPOL 179
Db 142 YPGLRGSPARDWTFVFLLEELLQAFKNEAQLTMRPRLLSAAVSGDPHVIQKAYDARLL 201
QY 180 SQYLDYIHWMTYDLHGSWEGYTGENSEPLYKYPTDTGNSNAYLVNVDYVMYWKDNGAPAEKL 239
Db 202 GRLDFISVLSYDLHGSWEKVTGHSNPLSLPGLDPPKSSA-----YAMSVMROLGVPPEKL 256
QY 240 IVGFPTYGHNFILSNPSNTGIGAPTSAGAGPAGPYAKESGIWAYEICTFLKNGATQGWDA 299
Db 257 LMGFTYGTFTHLRASNQELGAGAAGPASPCKYKQAGFLAYEVCVSFVOR-AKRWIN 315
QY 300 PQEVYAYQGNVWYDNTKSFEDIKAQWLKHNKFGGAMVWAIDLDFTCNQGKFFPLI 359
Db 316 DQYVPYAFKGEWVGYYDAISFGYKAFKIKRHFHGGAMVWTLDDDFRGNFGCTGPFPLA 375
QY 360 STLKAL 366
Db 376 HTLNLL 382

RESULT 9
OGP_PIG STANDARD; PRT; 527 AA.
AC Q28990;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Oviduct-specific glycoprotein precursor (Oviductal glycoprotein)
DE (oviductin) (estrogen-dependent oviduct protein) (POSP-E3).
GN OVGPI OR OGP.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=Yorkshire X Duroc X Hampshire; TISSUE=Oviduct;
RA MEDLINE=97107140; PubMed=894988;
RX Buhi W.C., Alvarez I.M., Choi I., Cleaver B.D., Simmen F.A.;
RT "Molecular cloning and characterization of an estrogen-dependent
porcine oviductal secretory glycoprotein."
RL Biol. Reprod. 55:1305-1314(1996).
CC -!- FUNCTION: BINDS TO OOCYTE ZONA PELLUCIDA IN VIVO. MAY PLAY A ROLE
IN THE FERTILIZATION PROCESS AND/OR EARLY EMBRYONIC DEVELOPMENT.
CC -!- SUBCELLULAR LOCATION: Secretory granules.
CC -!- TISSUE SPECIFICITY: OVIDUCT.
CC -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U43490; AAA85445.1;
DR InterPro; IPR001579; Chitinase_18/2.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
KW Glycoprotein; Fertilization; Signal.
FT SIGNAL 1 21 BY SIMILARITY.
FT CHAIN 22 527 OVIDUCT-SPECIFIC GLYCOPROTEIN.
FT CARBOHYD 62 62 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 402 402 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 527 AA; 58519 MW; 3187849CA2363A2 CRC64;
Query Match 37.4%; Score 936.5; DB 1; Length 527;
Best Local Similarity 47.3%; Pred. No. 66-61;
Matches 186; Conservative 61; Mismatches 123; Indels 23; Gaps 5;
QY 1 YOLTCYFTNWQYRGLGRFMPDNIDPCLCTHLIYAFAGRONNEITTIENWDTL-YQAF 59
Db 22 HKLVCFYTNWAFSRPGSASILPRDLDPFLCTHLVFAFASMMDSQIVAKDARDESIFYPEF 81
QY 60 NGLKNKNSQLTKLLAIGWNGFTAPFTAMVSTPENROTEITSVIRKLOYEDGLDFDWE 119
Db 82 NKLKERNRGLTKLLSVGGWNGFTSRFTKMLSTFSNRERVKSVIALLRHFGDGLDLPL 141
QY 120 YPGSRGSPQDKHLFTVLVOEMREAFEOBAKQINPKRLMVTAAVAGISNIOGSVEIPOL 179
Db 142 YPGLRGSPARDWTFVFLLEELLQAFKNEAQLTMRPRLLSAAVSGDPHVIQKAYDARLL 201
QY 180 SQYLDYIHWMTYDLHGSWEGYTGENSEPLYKYPTDTGNSNAYLVNVDYVMYWKDNGAPAEKL 239
Db 202 GRLDFISVLSYDLHGSWEKVTGHSNPLSLPGLDPPKSSA-----YTMVYWKLGAPPEKL 256
QY 240 IVGFPTYGHNFILSNPSNTGIGAPTSAGAGPAGPYAKESGIWAYEICTFLKNGATQGWDA 299
Db 257 LMGFTYGTFTHLRASNQELGAGAAGPASPCKYKQAGFLAYEVCVSFVOR-AKRWID 315
QY 300 PQEVYAYQGNVWYDNTKSFEDIKAQWLKHNKFGGAMVWAIDLDFTCNQGKFFPLI 359
Db 316 DQYVPYAFKGEWVGYYDAISFGYKAFKIKRHFHGGAMVWTLDDDFRGNFGCTGPFPLV 375
QY 360 STLKAL 376
Db 376 YMLNLLKAEVSSITSLSPGFLSTTVNSRTPC 408

RESULT 10
CHIT CAEEL STANDARD; PRT; 617 AA.
ID CHIT CAEEL
AC Q11174;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Putative endochitinase (EC 3.2.1.14).
GN C04F6.3
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Nham M.;
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages of N-
CC acetyl-D-glucosamine polymers of chitin.
CC -!- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
CC HYDROLASES).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U42835; AAA83586.1;
DR WormPeP; C04F6.3; CE03923.
DR InterPro; IPR002557; Chitin_bind_Pera.
DR InterPro; IPR001579; Chitinase_18/2.
DR InterPro; IPR001223; Glyco_hydro_18.


```
DR SMART; SM00494; ChtBD2; 1.
DR PROSITE; PS01095; CHITINASE 18; 1.
KW Hydrolase; Glycosidase; Chitin degradation; Signal; Antigen;
KW Repeat; Glycoprotein; Calcium-binding.
FT SIGNAL 1 22
FT CHAIN 23 504
FT DOMAIN 23 400
FT DOMAIN 401 450
FT DOMAIN 407 448
FT ACT SITE 148 148
FT ACT SITE 148 148
SQ SEQUENCE 504 AA; 55971 MW; A78BE7BF8E3709B CR64;

Query Match 31.4%; Score 786.5; DB 1; Length 504;
Best Local Similarity 35.0%; Pred. No. 4.9e-50;
Matches 169; Conservative 73; Mismatches 200; Indels 41; Gaps 9;

QY 1 YQLTCYFTWAQYRPGCLGRPMEDNDPCCLTHLIYAFAG-RQNNETITIEWDVT-----54
DB 23 YRGCYTTWAQYRQDEGKFLPGNPGNGLCTHLYAFKVDLGDGSKPFEWDEDETSK 82
QY 55 -LYQAFNGLNKNSQLKTLAIGWNFGTAPFTAMVSTPENRQTFITSVIKFLRQVEFDG 113
DB 83 GMSYATKLRNTPGLKVLVSGYNGFGSAITGIAKSAQKTERFIKSAIFLRKNFPG 142
QY 114 LDFDWEYPSGRSPPODKHLFTVLVQEMREAFQEAQINKPRLMVTAAAGISNIQSG 173
DB 143 FDLWDEYVPG---VAEEH-AKLVEAMKTAFAVEAKTSGKQRLLLTAASVAGKGTIDGS 196
QY 174 YEIPQLSQVLDYHVMYTHLHSGWEGYGENSPLYKYPTDTSNAYLVNVDYVMYWKNG 233
DB 197 YNVEISLGRNFDLLFLMSYDLHSGWKNVDLHGKLPKTEGVSGIGIFNTEFAADYVASKG 256
QY 234 APAELIVGPTVGHNFILSNPNTGIGAPTSGAGPAGYAKESGIWAYYEICTFLKNGA 293
DB 257 MPEKIIIGIPYACQWTLNDSETAIGAAASRPSSAKTNPAGGTASTWEICKYLKEGG 316
QY 294 TQGDAPQEVYVAYQGNVWGVYDNIKSPDKIAQWLKHNKFGGAMWAIDLDTFTGFCNQ 353
DB 317 KETVHOGVGAYVWKGQWYGVYDNEETIRIKMKWLKEGKGAFIWLDFDFTGKSCGK 376
QY 354 GKFFLITLKKALGLOSASCTAPAQPIEP-ITAAPSGSGNGSGSSS-----398
DB 377 GPYPLNLAISSELEGESENPEITTE--EPSITETAYETDETEETSETAYDTDETEETS 434
QY 399 -----SGSGSGSGGFCVANGILYVANNRNFAFWHCVNGVYVQCNQAGLGFDTSCD 450
DB 435 ETEATYDTEDEGQE--CPEP-DGLFPHTDCHLFIQCANNIAYVMQCPATTFENDAIK 491
QY 451 CCN 453
DB 492 VCD 494

RESULT 14
CHIL_APHAL
ID CHIL_APHAL STANDARD; PRT; 423 AA.
AC P32470;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Chitinase 1 precursor (EC 3.2.1.14).
GN CHIL.
OS Aphanocladium album.
OC Eukaryota; Fungi; Ascomycota; mitosporic Ascomycota; Aphanocladium.
OX NCBI_TaxID:12942;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ETHM 483;
RA MEDLINE=93013040; PubMed=1398137;
RX Blaiseau P.-L., Lafay J.-F.;
RT "Primary structure of a chitinase-encoding gene (chil) from the filamentous fungus Aphanocladium album: similarity to bacterial chitinases.";
```

ID CH11_BACCI STANDARD; PRT; 699 AA.
AC P20533;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE Chitinase A1 precursor (EC 3.2.1.14).
GN CHIA1.
OS Bacillus circulans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1397;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WL-12;
RX MEDLINE=90368776; PubMed=2203782;
RA Watanabe T., Suzuki K., Oyanagi W., Ohnishi K., Tanaka H.;
RT "Gene cloning of chitinase A1 from Bacillus circulans WL-12 revealed
RT its evolutionary relationship to Serratia chitinase and to the type
RT III homology units of fibronectin.";
RL J. Biol. Chem. 265:15659-15665 (1990).
RN [2]
RP MUTAGENESIS.
RC STRAIN=WL-12;
RX MEDLINE=93366760; PubMed=8103047;
RA Watanabe T., Kohori K., Miyashita K., Fujii T., Sakai H.,
RA Uchida M., Tanaka H.;
RT "Identification of glutamic acid 204 and aspartic acid 200 in
RT chitinase A1 of Bacillus circulans WL-12 as essential residues for
RT chitinase activity.";
RL J. Biol. Chem. 268:18567-18572 (1993).
CC -!- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages of N-
CC acetyl-D-glucosamine polymers of chitin.
CC -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC -!- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
CC HYDROLASES).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M57601; AAA81528.1; -.
CC PIR; A38368; A38368.
CC HSP; P02751; LFNA.
CC InterPro; IPR003610; CBM_5.12.
CC InterPro; IPR001579; Chitinase_18/2.
CC InterPro; IPR003961; FN_III.
CC InterPro; IPR003962; FNIII repeat.
CC InterPro; IPR001223; Glyco_hydro_18.
CC Pfam; PF00041; fn3; 2.
CC Pfam; PF00704; Glyco_hydro_18; 1.
CC Pfam; PF02839; CBM_5_12; 1.
CC PRINTS; PR00014; FNTYPEIII.
CC ProDom; PD000471; Glyco_hydro_18; 2.
CC SMART; SM00495; ChtBD3; 1.
CC SMART; SM00060; FN3; 2.
CC PROSITE; PS01095; CHITINASE_18; 1.
KW Hydrolase; Glycosidase; Chitin degradation; Signal; Repeat.
FT SIGNAL 1 41
FT CHAIN 42 699
FT DOMAIN 42 460
FT CATALYTIC
FT FIBRONECTIN TYPE-III (R-1).
FT FIBRONECTIN TYPE-III (R-2).
FT PROTON DONOR (PROBABLE).
FT D->N: DECREASE IN ACTIVITY.
FT D->E: NO CHANGE IN ACTIVITY.
FT E->D,Q: LOSS OF ACTIVITY.
FT MUTAGEN 204 204
FT MUTAGEN 699 AA; 73677 MW; ACTC9B22E2987643 CRC64;
SQ SEQUENCE

Query Match

Best Local Similarity 18.5%; Score 463.5; DB 1; Length 699;

Pred. No. 2.6e-26;

Matches 129; Conservative 76; Mismatches 172; Indels 85; Gaps 17;
QY 1 YQLTCYFTNWAQYRPLGR-FMPDNIDPCLCTHLIYAFAG-----SPP 39
DB 44 YKIVGYFSWAAY-----GRYNNVADIDFTKTHINAFADICWNGIHGPNPDSGPNPTW 99
QY 40 -RQNEITITIE-----WNDVTLYQA-----FNLGNKNSQLKTL 73
DB 100 TCQNEKSTINVPNGTIVLGDPIWDTGKTFTAGDTWDQPIAGNINQLKTKQTPNPKTII 159
QY 74 AIGWNFGTAPTAMVSTPENRQTPIITSVIFELROYEEGLDFDWEYPCSRG-----SPP 128
DB 160 SVGGWTWSNR-FSDVAATAATREVFANSAVDPLRKYNFDGVDLDWEYVPSGGLDGNKRKP 218
QY 129 ODKHLFTVLVOEMREAFQEAQINKPRLMTVAAGAAGISNIQSGYEIPQLSOYLDYIHV 188
DB 219 EDKQNYTILLISKIREKLDAGAVDOKKYLITIASGAS--ATYAANTEAKIAIIVDWINI 276
QY 189 MTYDLHGSWEGYTGENSEPLYKYPTDTGSNA-----YLNVDYVNNYWKDNGAPAEKLIIVGPP 244
DB 277 MTYDPNGAWQKISAHNAPLNVDPAASAAGVDPDANTFNVAAGAQGHLDAGVPAKLVLGVP 336
QY 245 TYGHNFILSNPNTGIGAPTSAGAGPAGYAKESGIWAYYEI-CTFL-KNGATQGHDAPOE 302
DB 337 FYGRGWDGCAQAGNGQYQTCCTGSSVGTW--EAGSFDFYDLEANYINKNGYTRYWNTAK 394
QY 303 VPYAYQGN--VMVGYDNITKSPDKAOWLKHNFKFGAMVWAIDDDFTGTFCNOGKFPPLIS 360
DB 395 VPYLYNASKRFISYDDAESGVYKTAYIKSKGLGGMFWELS-GDRNKTQLQKLLKADL-- 451
QY 361 TLKKALGLQASCTAPAPQPIBPITAAPSGSGNGSGSSSSSGS 402
DB 452 -----PTGGTVP--PVD--TTAPSVPGNARSTGVTANS 480

Search completed: June 29, 2003, 21:00:23

Job time : 14.0769 secs

THIS PAGE BLANK (USPTO)

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 29, 2003, 21:10:28 ; Search time 1499.62 Seconds
(without alignments)
4913.891 Million cell updates/sec

Title: US-10-004-219b-14

Perfect score: 2506

Sequence: 1 YQLTCYFTNQAQYRPLGRF.....QQNCOAGLVFTSDCCNWA 455

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DEV=xlp

-Q=/cgn2_1/USPTO_spool/US10004219/runat_24062003_160228_8173/app.query.fasta_1.1877

-DB=EST -QFMT=FASTAP -SUFFIX=first -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0

-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45

-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL

-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000

-USER=US10004219 @CGN 1 1.4575 @runat_24062003_160228_8173 -NCPU=6 -ICPU=3

-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG

-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6

-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: em_estba.*
- 2: em_esthum.*
- 3: em_estin.*
- 4: em_estmu.*
- 5: em_estov.*
- 6: em_estpl.*
- 7: em_estro.*
- 8: em_hcc.*
- 9: gb_est1.*
- 10: gb_est2.*
- 11: gb_hcc.*
- 12: gb_est3.*
- 13: gb_est4.*
- 14: gb_est5.*
- 15: em_estfun.*
- 16: em_estom.*
- 17: gb_gss.*
- 18: em_gss_hum.*
- 19: em_gss_inv.*
- 20: em_gss_pnt.*
- 21: em_gss_vrt.*
- 22: em_gss_fun.*
- 23: em_gss_nam.*
- 24: em_gss_mus.*
- 25: em_gss_other.*
- 26: em_gss_pro.*
- 27: em_gss_rtd.*

ALIGNMENTS

RESULT 1 AK008633 1533 bp mRNA linear HTC 19-JAN-2002

AK008633 Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:2200003E03:chitinase, acidic, full insert sequence.

DEFINITION AK008633 GI:12842941

ACCESSION AK008633

VERSION HTC; CAP trapper.

KEYWORDS Mus musculus (strain:C57BL/6J) adult male stomach cDNA to mRNA,

SOURCE clone lib:RIKEN full-length enriched mouse cDNA library

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1

Result No.	Score	Query Match	Length	DB	ID	Description
1	2097.5	83.7	1533	11	AK008633	AK008633 Mus muscu
2	2079.5	83.0	1535	11	AK008757	AK008757 Mus muscu
3	2025.5	80.8	1534	11	AK008650	AK008650 Mus muscu
4	1980	79.0	1530	11	AK007573	AK007573 Mus muscu
5	1438.5	57.4	957	13	BI760250	BI760250 603045190
6	1375	54.9	1105	14	BM923480	BM923480 AGENCOURT
7	1327.5	53.0	892	14	BQ231098	BQ231098 AGENCOURT
8	1270	50.7	838	12	BG867815	BG867815 602786336
9	1251	49.9	1038	13	BI517797	BI517797 603042049
10	1241	49.5	872	14	BQ231064	BQ231064 AGENCOURT
11	1237	49.4	809	12	BG866417	BG866417 602785352
12	1209	48.2	764	12	BG869304	BG869304 602788819
13	1207	48.2	763	12	BG872647	BG872647 602793705
14	1205	48.1	837	12	BG868741	BG868741 602787570
15	1202	48.0	939	14	BQ219632	BQ219632 AGENCOURT
16	1201	47.9	857	12	BG871434	BG871434 602790636
17	1198	47.8	894	12	BG869958	BG869958 602789826
18	1195.5	47.7	843	12	BG868207	BG868207 602784727
19	1194	47.6	1100	12	BF302298	BF302298 602031269
20	1188.5	47.4	855	12	BG872136	BG872136 602792974
21	1182	47.2	833	12	BG871337	BG871337 602790509
22	1181.5	47.1	949	12	BG868019	BG868019 602788381
23	1180.5	47.1	886	12	BG870840	BG870840 602791891
24	1178	47.0	824	13	BI518221	BI518221 603042049
25	1178	47.0	880	12	BG871468	BG871468 602790678
26	1172	46.8	847	12	BG866775	BG866775 602786605
27	1170.5	46.7	848	12	BG869143	BG869143 602789920
28	1170	46.7	856	12	BG870147	BG870147 602793390
29	1169	46.6	1155	12	BG867473	BG867473 602788716
30	1167.5	46.6	836	12	BG869655	BG869655 602789716
31	1167.5	46.6	976	14	BQ225554	BQ225554 AGENCOURT
32	1167	46.6	800	12	BG866853	BG866853 602786689
33	1164.5	46.5	788	12	BG873335	BG873335 602794304
34	1162.5	46.4	900	12	BG867646	BG867646 602787435
35	1161	46.3	830	12	BG869896	BG869896 602789435
36	1160	46.3	923	13	BI686954	BI686954 603313831
37	1159.5	46.3	882	12	BG873488	BG873488 602791730
38	1159	46.2	742	12	BG873143	BG873143 602794547
39	1157.5	46.2	863	12	BG872850	BG872850 602793373
40	1156.5	46.1	874	12	BG865756	BG865756 602784345
41	1156.5	46.1	948	13	BI664452	BI664452 603290028
42	1155	46.1	828	12	BG869811	BG869811 602789327
43	1151.5	45.9	845	12	BG865848	BG865848 602788264
44	1150	45.9	757	12	BG867292	BG867292 602788635
45	1150	45.9	837	12	BG865578	BG865578 602783726

Db 260 GACTTGAATAAAGCAGCAACAGCAAACTGAAACCCCTCTCTGGCAATTTGGAGGCTTGAACCTTT 319
Qy 81 GlyThrAlaProPheThrAlaMetValSerThrProGluAenArgGlnThrPheIleThr 100
Db 320 GGAAGCTGCTCTTTTCACTACCATGTTTCCACTTCTCAGAACCCGACAGACCTTCATTATACC 379
Qy 101 SerValIleLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTyrGluTyr 120
Db 380 TCAGTCATCAAAATTTCTGGTCAGTATGGTTTGTGACCTGGACCTGGAGCTGGGAATAC 439
Qy 121 ProGlySerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValGlnGlu 140
Db 440 CCAGGCTCAGCTGGAGCCCTCTCAGCAACAGCATCTCTTCACTGTCTCTGTGTGAAGAA 499
Qy 141 MetArgGluAlaPheGluGlnGluAlaLysGlnIleAenLysProArgLeuMetValThr 160
Db 500 ATGCGTGAAGCTTTTGGAGGAGGCTATTGAGAGCAACAGGCCCGACGACTGATGTTTACT 559
Qy 161 AlaAlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSer 180
Db 560 GCTGCTGTAGCTGTGGGATTTTCAACATCCAGGCTGGCTATGAGATCCCTGAACCTTCT 619
Qy 181 GlnTyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTyrGluGlyTyr 200
Db 620 AAGTACCTGGATTTTCACTCATGATCATATGATGATGATGATGATGATGATGATGATGAT 679
Qy 201 ThrGlyGluAenSerProLeuTyrLysTyrProThrAspThrGlySerAsnAlaTyrLeu 220
Db 680 ACTGGGGAGATAGTCTCTTTTACAAATACCTCTAGTACGATGGTGAAGATCCCTACCTC 739
Qy 221 AsnValAspTyrValMetAsnTyrTyrLysAspAsnGlyAlaProAlaGluLysLeuIle 240
Db 740 AATGTGGATTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 799
Qy 241 ValGlyPheProThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIle 260
Db 800 GTTGGATTTCCAGAGATATGAGACACACCTTCATCTCTGAGAAACCCCTCTCATATAATGGAAT 859
Qy 261 GlyAlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyIleTyr 280
Db 860 GTGTCCTCTTACCTCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 919
Qy 281 AlaTyrTyrGluIleCysThrPheLeuLysAsnGlyAlaThrGlnGlyTyrAspAlaPro 300
Db 920 GCCTACTATGATTTGACCTTTCTGAGAGTGGAGCCACTGAGGTCTGGGATGCTCTCC 979
Qy 301 GlnGluValProTyrAlaTyrGlnGlyAenValTyrValGlyTyrAspAsnIleLysSer 320
Db 980 CAAGAAGTGCCTATGCTTATAGCCCAACGAGTGGCTTGGCTATGACAAATATCAAGAGC 1039
Qy 321 PheAspIleLysAlaGlnTyrLeuLysHisAsnLysPheGlyGlyAlaMetValTyrAla 340
Db 1040 TTCAGTGTAAAGGCTCAGTGGCTTAAAGCAGAAACAATTTTGGAGGTGCCATGATCTGGGCC 1099
Qy 341 IleAspLeuAspAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSer 360
Db 1100 ATTGACCTTGATGACTTCACTGGCTTCTCTGTGATCAGGGAATTTCTCTGACTTCT 1159
Qy 361 ThrLeuLysLysAlaLeuGlyLeuGlnSerAlaSerCysThrAlaProAlaGlnProIle 380
Db 1160 ACTTTTGAACAAAGCCCTTGGCATATCACTGAAGGTGGCAGAGTCTGACGCTGCTCTCC 1219
Qy 381 GluProIleThrAlaAlaProSerGlySerGlyAsnGlySerGlySerSerSerGly 400
Db 1220 GAGCCAGTGAATCTCTCTCA-----GGAAGTGGAGGTGGGGGTGAAGCTCCGGA 1270
Qy 401 GlySerSerGlySerGlyPheCysAlaValArgAlaAsnGlyLeuTyrProValAla 420
Db 1271 GGAAGCTCTGGAGGAGTGGATTTCTGTGCCGCAAGCAGATGGCTCTACCTCTGTGGCA 1330
Qy 421 AsnAenArgAsnAlaPheThrHisCysValAenGlyValThrTyrGlnAenCysGln 440
Db 440

Db 1331 GATGACAGAAATGCTTTTGGCAGTCATCAATGGAATCATCATACCAGCAGCATGTGCAA 1390
Qy 441 AlaGlyLeuValPheAspThrSerCysAspCysCysAsnTyr 454
Db 1391 GCAGGGCTTGTTTTGTATACCATGCTGTAATTTGCTGCAACTGG 1432
RESULT 2
AK008757 1535 bp mRNA linear HTC 19-JAN-2002
LOCUS Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:2210020N23:chitinase, acidic, full insert sequence.
DEFINITION AK008757 GI:12843148
ACCESSION AK008757
VERSION Mus musculus (strain:C57BL/6J) adult male stomach cDNA to mRNA,
KEYWORDS clone lib:RIKEN full-length enriched mouse cDNA library
SOURCE clone:2210020N23.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Carninci, P. and Hayashizaki, Y.
2
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636
3
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159
4
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yanamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Orawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
5
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,
Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
Fleischmann, W., Gaasterland, T., Glissi, C., King, B., Kochiwa, H.,
Kuhl, P., Lewis, S., Matsuo, Y., Nikaudo, I., Pesole, G.,
Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M.,
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
Carninci, P., de Bonaldo, M. F., Brownstein, M. J., But, C.,
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P.,
Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,
Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H.,
Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L.,
Yoshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kotsuki, S.
and Hayashizaki, Y.
Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
21085660
11217851
5 (bases 1 to 1535)
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,

Qy 343 LeuAspAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSerThrIleu 362
 Db 1106 CTTGATGACTTCACTGGCTCTTCTGTGATCAGGAGAAATTTCTCTGACTTCTACTTTG 1165
 Qy 363 LysLysAlaLeuGlyLeuGlnSerLysAlaSerCysThrAlaProAlaGlnProIleGlnPro 382
 Db 1166 AACAAAGCCCTTGGCATATCCATCAAGGTTGCACAGCTCTCGAGTCCGAGGCA 1225
 Qy 383 IleThrAlaAlaProSerGlySerGlyAsnGlySerGlySerSerSerSerGlyGlySer 402
 Db 1226 GTGACTACTCTCCA-----GGAAGTGGAGTGGGGTGGAGTCCGAGGAGAGC 1276
 Qy 403 SerGlyGlySerGlyPheCysAlaValArgAlaAsnGlyLeuTyProValAlaAsn 422
 Db 1277 TCTGGAGGAGTGGATTCTGTGCCACAAAGCAGATGGCTCTACCTGTGGCAGATGAC 1336
 Qy 423 ArgAsnAlaPheThrPheCysValAsnGlyValThrTyGlnAlaAsnCysGlnAlaGly 442
 Db 1337 AGAAATGCTTTTGGCAGTGCATCAATGGAATCATACACAGCAGCATTTGTCAGCAGG 1396
 Qy 443 LeuValPheAspThrSerCysAspCysCysAsnTtp 454
 Db 1397 CTGTGTTTGTATACAGCTGTAATTTGCTGCACTGG 1432
 RESULT 3
 Locus AK008650
 DEFINITION Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:2200008L17:chitinase, acidic, full insert sequenced.
 ACCESSION AK008650
 VERSION AK008650.1 GI:12842971
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (strain: C57BL/6J) adult male stomach cDNA to mRNA, clone lib:RIKEN full-length enriched mouse cDNA library
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1
 AUTHORS Carninci, P. and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636
 REFERENCE 2
 AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE 20499374
 PUBMED 11042159
 REFERENCE 3
 AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kaishwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Onoda, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipipillary sequencer.
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
 MEDLINE 20530913
 PUBMED 11076861
 REFERENCE 4
 AUTHORS Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I., Saito, T., Okazaki, Y., Gojohori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,

Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Balzoni, L., Mashima, J., Mazzarelli, J., Sakamoto, N., Sasaki, H., Ring, B., Ringwald, C., Seva, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wyshaw-Borish, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kotsuki, S. and Hayashizaki, Y.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409 (6821), 685-690 (2001)
 PUBMED 21085660
 11217851
 5 (bases 1 to 1534)
 Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Hori, F., Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J., Schriml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamana, I., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
 Direct Submission
 Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Kanagawa 230-0045, Japan [E-mail: genome-res@gs.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216]
 Please visit our web site. (http://genome.gsc.riken.go.jp/) for further details.
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'-GAGAGAGAGAGATCCAGAGCTCTTTTITTTTITVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5'-GAGAGAGATTCCTCGATTAAATTAATTAATCCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.
 Host: SOUR.

FEATURES
 Source
 1. 1534
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="FANTOM_DB:2200008L17"
 /db_xref="MGD:MG1:1894268"
 /db_xref="taxon:10090"
 /clone="2200008L17"
 /sex="male"
 /tissue_type="stomach"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="adult"
 1. 1534
 /note="Chitinase, acidic"
 data source:MGD, source key:MG1:1932052, evidence:ISS"
 /db_xref="MGD:MG1:1932052"
 misc_feature
 378 a 383 c 374 g 399 t
 BASE COUNT
 ORIGIN

Alignment Scores:

Pred. No.: 1.32e-204 Length: 1534
 Score: 2025.50 Matches: 366
 Percent Similarity: 89.43% Conservative: 40
 Best Local Similarity: 80.62% Mismatches: 45
 Query Match: 80.83% Indels: 5
 DB: 11 Gaps: 1

US-10-004-219b-14 (1-455) x AK008650 (1-1534)

QY 1 TyrGlnLeuThrCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPhe 20
 DB 80 TACAATCTCATATGCTATTTTCAACCACTGGGCCAGATATCGGCCAGGCTGGGGAGCTTC 139
 QY 21 MetProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArg 40
 DB 140 AAGCCTGATGACATTAACCCCTGCTGTGTACTCACCTGATCTATGCCCTTGTGGGATG 199
 QY 41 GlnAsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrGlnAlaPheAsn 60
 DB 200 CAGAACATGAGATACACCACTAGATGAATGATGTTTACTCTCTATAAAGCTTTCAT 259
 QY 61 GlyLeuLeuAsnIleAsnSerGlnLeuLeuThrHisLeuAlaIleGlyGlyTrpAsnPhe 80
 DB 260 GACTTTGAAAAACAGGAACAGCAACCTCGGCCAATGGAGGCTGGAACTTT 319
 QY 81 GlyThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThr 100
 DB 320 GGAACCTGCTCTTCACTACCATGGTTTCCACTTCTCAGAACCGCCAGACCTTCATTACC 379
 QY 101 SerValIleLeuPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyr 120
 DB 380 TCAGTCATCAAAATTTCTGCTCAGTATGGTTTGTATGGAGCTGACCTGGGAATAC 439
 QY 121 ProGlySerArgGlySerProProGlnAspLeuHisLeuPheThrValLeuValGlnGlu 140
 DB 440 CCAGGCTCAGTGGGAGCCCTCTCCAGCAACAGCATCTCTTCACTGTCTCGTGGAGGAA 499
 QY 141 MetArgGluAlaPheGluGlnGluAlaLeuGlnIleAsnLeuProArgLeuMetValThr 160
 DB 500 ATGCTGGAAGCTTTTGAGCAGGAGGCTATTGAGCAACAGCCCAAGCTGATGTTACT 559
 QY 161 AlaAlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSer 180
 DB 560 GCTGCTGTAGTGGTGGGATTTCCAACTCCAGGCTGGCTATGAGATCCCTGAACCTTCT 619
 QY 181 GlnTyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr 200
 DB 620 AAGTACCTGGATTTTCATCTCATGTATGATGACCTCATGGCTTCTGGGAGGCTAC 679
 QY 201 ThrGlyGluAsnSerProLeuTyrIleTyrProThrAspThrGlySerAsnAlaTyrLeu 220
 DB 680 ACTGGGGA-AAATGCTCTTTTACAATACCTTACTGAGACTGGTAGCAATGCCTACCTC 738
 QY 221 AsnValAspTyrValMetAsnTyrTrpIleAspAsnGlyAlaProAlaGluLeuLeu 240
 DB 739 AATGTGATATTATGTCATGAATATTGGAAGAACATGAGGCCCACTGAGAAGCTCAT 798
 QY 241 ValGlyPheProThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIle 260
 DB 799 GTTGGATC-CGAGAGPATGCAACACCTTCTATCTGAGAAACCCCTCTCATATGGAAT 857
 QY 261 GlyAlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaIleGluSerGlyIleTrp 280
 DB 858 GTTGGCCCTACTCTGCTGATGGCCCTGCTGGCCCTATACAGACAGGCTGGGTCTGG 917
 QY 281 AlaTyrTyrGluIleCysThrPheLeuLeuAsnGlyAlaThrGlnGlyTrpAspAlaPro 300
 DB 918 GCTACTATGAGATTGCACTTCTTGAAGAGTGGAGCCCACTGAGGTCTGGATGCTCC 977
 QY 301 GlnGluValProTyrAlaTyrGlnGlyAsnValTrpValGlyTyrAspAsnIleYsSer 320
 DB 978 CAAGAAGTCCCTATGCCCTATAAGCAACGAGTGGCTTGGCTATGACAATATCAAGAGC 1037

QY 321 PheAspIleLeuAlaGlnTrpLeuIleHisAsnLysPheGlyGlyValaMetValTrpAla 340
 DB 1038 TTCAGTGTTAAGGCTCAGTGGCTTAAGCAGAACAAATTTTGGAGGTGCCATGATCTGGGCC 1097
 QY 341 IleAspLeuAspAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSer 360
 DB 1098 ATTGACCTTGATGACTTCACTGGCTCTTCTGTGATCAGGGAATAATTTCTCTGACTTCT 1157
 QY 361 ThrLeuLysLeuAlaLeuGlyLeuGlnSerAlaSerCysThrAlaProAlaGlnProIle 380
 DB 1158 ACTTTGAACAAAGCCCTTGGCATATCCACTGAAGTTGCAGACTCCTGACGTGCTTCC 1217
 QY 381 GluProIleThrAlaAlaProSerGlySerGlyAsnGlySerGlySerSerSerGly 400
 DB 1218 GAGCCAGTGACTACTCTCTCCA-----GGAAGTGGAGTGGGGGTGGAAGCTCCGA 1268
 QY 401 GlySerSerGlyGlySerGlyPheCysAlaValArgAlaAsnGlyLeuTyrProValAla 420
 DB 1269 GGAAGCTCTGGAGGAGTGGATTTCTGTCCGCAACAAAGCAGATGGCTCTTACCTGTGGCA 1328
 QY 421 AsnAsnArgAsnAlaPheThrHisCysValaAsnGlyValThrTyrGlnGlnAsnCysGln 440
 DB 1329 GATGACAGAAATGCTTTTGGCAGTGCATCAATGAATCACATACAGCAGCATTTGTCAA 1388
 QY 441 AlaGlyLeuValPheAspThrSerCysAspCysCysAsnTrp 454
 DB 1389 GCAGGCGCTGTTTGTATACCAAGCTGTAATTTCTGCTCAACTGG 1430
 RESULT 4
 AK007573 1530 bp mRNA linear HTC 19-JAN-2002
 LOCUS AK007573.1 GI:12841201
 DEFINITION Mus musculus 10 day old male pancreas cDNA, RIKEN full-length enriched library, clone:1810021D23:chitinase, acidic, full insert sequence.
 ACCESSION AK007573
 VERSION AK007573.1
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (strain:CS7BL/6J) 10 day old male pancreas cDNA to mRNA, clone lib-RIKEN full-length enriched mouse cDNA library
 clone:1810021D23.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 Carninci, P. and Hayashizaki, Y.
 AUTHORS High-efficiency full-length cDNA cloning
 TITLE Meth. Enzymol. 303, 19-44 (1999)
 JOURNAL 99279253
 MEDLINE 10349636
 PUBMED 10349636
 REFERENCE 2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 AUTHORS Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 TITLE Genome Res. 10 (10), 1617-1630 (2000)
 JOURNAL 20499374
 MEDLINE 11042159
 PUBMED 11042159
 REFERENCE 3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 AUTHORS RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
 TITLE Genome Res. 10 (11), 1757-1771 (2000)
 JOURNAL 20530913
 MEDLINE 11076861
 PUBMED 11076861
 REFERENCE 4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,


```
|||||
915 GCCTACTAGATTGACCTTTCTGAGAAGTGGAGCCACTGAGCTGGTGGATGCTCC 974
QY GlnGluValProTyrAlaTyrGlnGlyAsnValTyrValGlyTyrAspAsnIleLysSer 320
DB CAAGAAGTGGCTTATGCTATATAGGCAACAGAGTGGCTTGGCTATGACAAATATCAAGAGC 1034
QY PheAspIleLysAlaGlnTyrLeuLysHisAsnLysPheGlyGlyAlaMetValTyrAla 340
DB TTTCAGTGTAAAGCTCAGTGGCTTATAGCAGAACAAATTTGGAGTGCCATGATCTGGGC 1094
QY IleAspLeuAspAspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSer 360
DB ATTGACCTTGGATCACTGCTCTTCTGTGATCAGGGAATTTCTCTGACTTCT 1154
QY ThrLeuLysLysAlaLeuGlyLeuGlnSerAlaSerCysThrAlaProAlaGlnProIle 380
DB ACTTTGAACAAAGCCCTTGGCATATCCATGAAGTTGCACACCTCTGACGTGGCTTCC 1214
QY GluProIleThrAlaAlaProSerGlySerGlyAsnGlySerGlySerSerSerGly 400
DB GAGCCAGTACTACTCTCTCA-----GGAGTGGGAGTGGGGTGGAGCTCCGA 1265
QY GlySerSerGlyGlySerGlyPheCysAlaValArgAlaAsnGlyLeuTyrProValAla 420
DB GGAAGCTCTGGAGCGAGTGGATCTGTGCCGACAAAGCAGATGGCTTACCTGTGGCA 1325
QY AsnAsnArgAsnAlaPheThrHisCysValAsnGlyValThrTyrGlnGlnAsnCysGln 440
DB GATGACGAAATGCTTTTGGCAGTGCATATGGAATCACAATACAGCAGCATTTGCAA 1385
QY AlaGlyLeuValPheAspThrSerCysAspCysCysAsnTyr 454
DB GCAGGCTGTGTTTATACCACTGATATTTGTCGCACTGG 1427

RESULT 5
BI760250
LOCUS BI760250
DEFINITION 603045190F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5185486 5',
mRNA sequence.
ACCESSION BI760250
VERSION BI760250.1 GI:15751828
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 957)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAMI1463 row: c column: 23
High quality sequence stop: 780.
Location/Qualifiers
1..957
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5185486"
/clone_lib="NIH MGC_116"
/lab_host="DH10B"
/note="Organ: pooled colon, kidney, stomach; Vector:
pCMV-SPORT6; Site1: NotI; Site2: EcoRV (destroyed); RNA
source anonymous pool of 3 colons, age 26 yo male, 49 yo
female, 71 yo male colon; 46 yo male kidney, and pool of 2
```

FEATURES
Source

stomachs, 62 yo male and 70 yo female. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 023. Note: this is a NIH MGC Library."

BASE COUNT 222 a 277 c 227 g 231 t
ORIGIN

Alignment Scores:
Pred. No.: 2,05e-142 Length: 957
Score: 1438.50 Matches: 284
Percent Similarity: 94.16% Conservative: 6
Best Local Similarity: 92.21% Mismatches: 11
Query Match: 57.40% Indels: 8
DB: 13 Gaps: 1

US-10-004-219B-14 (1-455) x BI760250 (1-957)

QY 65 LysAsnSerGlnLeuLysThrLeuLeuAlaIleGlyGlyTyrAsnPheGlyThrAlaPro 84
DB 45 CGAACAGCCAGCTGAAAACTCTCTGGCCATGGAGGCTGGAATTCGGGACTGCCCT 104
QY 85 PheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThrSerValIleLys 104
DB 105 TTCACCTGCCATGTTTCTACTCTCTGAGAACGCCAGACTTTTCATCAGCTCAGTCATCAA 164
QY 105 PheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTyrGluTyrProGlySerArg 124
DB 165 TTCTCTGGCGCAGTATGAGTTTGACGGGCTGGACTTTGACTGGGAGTACCTGGCTCTCGT 224
QY 125 GlySerProGlnAspLysHisLeuPheThrValLeuValGlnGluMetArgGluAla 144
DB 225 GGGAGCCCTCTCCAGACAAAGCATCTCTTCACTGCTGTCGAGAAATGCTGAGCT 284
QY 145 PheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThrAlaAlaValAla 164
DB 285 TTTGAGCAGGAGGCCCAAGCAGATCAACAGCCAGCTGATGTCAGTGTAGCT 344
QY 165 AlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSerGlnTyrLeuAsp 184
DB 345 GCTGGCATCTCCAATATCCAGTCTGGCTATGAGATCCCCCACTGTCACAGTACCTGGAC 404
QY 185 TyrIleHisValMetThrTyrAspLeuHisGlySerTyrGluGlyTyrThrGlyGluAsn 204
DB 405 TACATCCATGTCATGACCTACGACCTCCATGCTCTGGAGGGCTTACTGAGAGAAC 464
QY 205 SerProLeuTyrLysTyrProThrAspThrGlySerAsnAlaTyrLeuAsnValAspTyr 224
DB 465 AGCCCCCTCTCAAAATACCCGACTGACACCGGCAGCAAGCCCTACTCAATGTGATTAT 524
QY 225 ValMetAsnTyrTyrLysAspAsnGlyAlaProAlaGluLysLeuIleValGlyPhePro 244
DB 525 GTTCATGAATCTTGGAGGACAAATGAGCAGCTGAGAGCTCATGCTGGATTCCT 584
QY 245 ThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIleGlyAlaProThr 264
DB 585 ACCTATGACACCAACTTATCTCTGAGCAACCCCTCCACACTGGAAATGGTGGCCCCCACC 644
QY 265 SerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyIleTyrAlaTyrTyrGlu 284
DB 645 TCTGTGTCTGTCTCTGTCGGCCCTATGCCAAGGAGTCTGGGATCTGGGCTTACTACGAG 704
QY 285 IleCysThrPheLeuLys-AsnGlyAlaThrGlnGlyTyrAspAlaProGln-GluValP 304
DB 705 ATCTGTACTCTTCTCAACAAATGAGGACCACTCAGGAGTGGGATGCCCTCAAGGAGTGC 764
QY 304 TGTyrAlaTyr-GlnGlyAsnValTyrValGlyTyr---AspAsnIleLysSerPheAsp 322
DB 765 CTTATGCCCTATTTCAGGCGCAGCTGTGGGTGGCTTATGACCAACATCAAGAGCTTTCAT 824
QY 323 IleLysAlaGlnTyrLeu-LysHisAsnLysPheGlyGlyAlaMetValTyrAlaIleAs 342

```

825 ATTAAGGCTCAATGGCTTTAGGCACACCAATTTGGAGGCGCATGGTCTCGCCATTGA 884
|||||
342 pleu-AspAspPhe-ThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSerThr 361
|||||
895 TCTCGATAACTTCCACTGGGCACTTCTGCAACCA-GGCAGTTTTTCCCTAATTTCCACC 943
|||||
362 LeuLysLysAla 365
|||||
944 TGTAAACCGCT 955
|||||

RESULT 6
BM923480 1105 bp mRNA linear EST 12-MAR-2002
LOCUS AGENCOURT 6623879 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5759117
DEFINITION 5', mRNA sequence.
ACCESSION BM923480
VERSION BM923480.1 GI:19373859
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1105)
REFERENCE NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgsabbs@mail.nih.gov.
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM12804 row: a column: 06
High quality sequence stop: 668.
Location/Qualifiers
1..1105
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5759117"
/clone_lib="NIH_MGC_116"
/lab_host="DH10B"
/note="Organ: pooled colon, kidney, stomach; Vector:
pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 3 colons, age 26 yo male, 49 yo
female, 71 yo male colon; 46 yo male kidney, and pool of 2
stomachs, 62 yo male and 70 yo female. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.4 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
023. Note: this is a NIH MGC Library."
BASE COUNT 270 a 318 c 285 g 231 t
ORIGIN

Alignment Scores:
Pred. No.: 1.49e-135 Length: 1105
Score: 1375.00 Matches: 280
Percent Similarity: 77.87% Conservative: 5
Best Local Similarity: 76.50% Mismatches: 18
Query Match: 54.87% Indels: 65
DB: 14 Gaps: 5

US-10-004-219B-14 (1-455) x BM923480 (1-1105)

65 LysAsnSerGlnLeuLysThrLeuLeuAlaIleGlyLysPheCysThrAlaPro 84
|||||
62 CGGACAGCCAGCTGAAACTCTCTGGGCAATGGAGGCTGGAACTTCGGGACTGC-CCC 120
|||||

```

RESULT 7

```

85 PheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThrSerValIleLys 104
120 -----
105 PheLeuArgGlnTyrGluPheAspClyLeuAspPheAspTyrGluTyrProGlySerArg 124
120 -----
125 GlySerProGlnAspLysHisLeuPheThrValLeuValGlnGlnMetArgGluAla 144
121 -----
145 PheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThrAlaAlaValAla 164
136 TTTGAGCAGGAGGGCCAGCAGATCAACAGCCAGGCTGATGCTCAGCTGCTGCAGTAGCT 195
165 AlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSerGlnTyrIleuAsp 184
196 GCTGGCATCTCAATATATCCAGTCTGGCTATGAGATCCCCCACTGTCACAGTACTGGAC 255
185 TyrIleHisValMetThrTyrAspLeuHisGlySerTyrGluGlyTyrThrGlyGluAsn 204
256 TACATCCATGTCTAGCCTACGACCTCCATGGCTCTGGAGGGGTACTACTGGAGAGAAC 315
205 SerProLeuTyrLysTyrProThrAspThrGlySerAsnAlaTyrLeuAsnValAspTyr 224
316 AGCCCCCTCTACAAATATCCGACTGACACCGCAGCAACGCTACTCAATGTGATTAT 375
225 ValMetAsnTyrTyrLysAspAsnGlyAlaProAlaGluLysLeuIleValGlyPhePro 244
376 GTCATGAATCTACTGGAGGCAATGGAGCACCAGCTGAGAAGCTCATCTGTGGATTTCCT 435
245 ThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIleGlyAlaProThr 264
436 ACCTATGGACACCACTTCATCTGTGACCAACCCCTCCACACCTGGATTGGTGGCCCCACC 495
265 SerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyIleTyrAlaTyrTyrGlu 284
496 TCTGGTGTCTGCTGCTGGGCCCTATGCCAAGGAGTCTGGGATCTGGGCTTACTACGAG 555
285 IleCysThrPheLeuLysAsnGlyAlaThrGlnGlyTyrAspAlaProGlnGluValPro 304
556 ATCTGTACTCTTCTGAAATAATGGAGCCACTCAGGAGTGGGATGGCCCTCAGGAAGTGCT 615
305 TyrAlaTyrGlnGlyAsnValTyrValGlyTyrAspAsnIleLysSerPheAspIleLys 324
616 TATGCTATTCAGGGCAATGTGTGGTTGGCTATGCAACATCAAGAGCTTCGATATTATAG 675
325 AlaGlnTyrLeuLysHisAsnLysPheGlyGlyAlaMetValTyrAlaIleAspLeuAsp 344
676 GCTCAATGGCTTAAGCACACAAATTTGGAGGCGCATGGTCTGGGCCATTGATCTGGAT 735
345 AspPheThrGlyThrPheCysAsnGlnGlyLysPheProLeuIleSerThr-Leu-LysL 364
736 GACTTCACTGGCACTTCTCTGCAC-CAGGGCAAGGTTCCCTTAATCTCCACCCCTGAAAGA 794
364 ysAlaLeuGlyLeuGlnSerAlaSerCysThr-AlaProAlaGlnProIleGluProIle 383
795 AGGCCCTCGGCTGCAGAGTGCAGTTGCACGGCTCCAGCTCAGCCCATTTAGGCAATA 854
384 ThrAlaAlaProSerGlySerGlyAsn---GlySerGlySerSerSer-----SerGly 400
855 ACTGTGCTCCAGTGGGAACGGGGAACCGGAGCGCGGAGTAGTAACAGCTCTCTGGAGGGC 914
401 GlySerSerGlyGlySerGlyPheCysAlaVal---ArgAlaAsnGly---LeuTyrPro 418
915 ACCTCGGGAAGCAATTGGAATTTCTGTGCTGTCTCAAAAGNCCACACGGGCCCTTCTTACCCC 974
419 ValAlaAsnAsn 422
975 GCTGGGGAATAAT 986

```

BQ231098 892 bp mRNA linear EST 02-MAY-2002
 LOCUS AGENCOURT 7578138 NCI_CGAP_St1 Mus musculus cDNA clone
 IMAGE:6051580 5', mRNA sequence.
 ACCESSION BQ231098
 VERSION BQ231098.1 GI:20412498
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 892)
 AUTHORS NIH-MGC http://imgc.nci.nih.gov/
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs@mail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM13305 row: k column: 05
 High quality sequence stop: 772.
 Location/Qualifiers
 1..892
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone_lib="IMAGE:6051580"
 /clone_lib="NCI CGAP St1"
 /lab_host="DH10B (T1-resistant)"
 /note="Organ: stomach; Vector: pCMV-SPORT6; Site 1: SalI;
 Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 1.77 Kb. Library constructed by Life
 Technologies. Note: this is a NCI CGAP Library."
 219 a 226 c 216 g 230 t 1 others

BASE COUNT 219 a 226 c 216 g 230 t 1 others

ORIGIN
 Alignment Scores:
 Pred. No.: 1,22e-130 Length: 892
 Score: 1327.50 Matches: 242
 Percent Similarity: 87.21% Conservative: 17
 Best Local Similarity: 81.48% Mismatches: 19
 Query Match: 52.97% Indels: 19
 DB: 14 Gaps: 1

US-10-004-219b-14 (1-455) x BQ231098 (1-892)

QY 1 TyrGlnLeuThrCysTyrPheThrAsnTpaAlaGlnTyrArgProGlyLeuGlyArgPhe 20
 Db 36 TACAATCTGATATGCTATTATTCACCACTGGGCCAGTATCGCCAGGTCTGGGAGCTTC 95
 QY 21 MetProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArg 40
 Db 96 AAGCCTGATGACATTAAACCCCTGCTGTACTACCTGATCTATGCTTGTGGGATG 155
 QY 41 GlnAsnAsnGluIleThrIleGluTpaAsnAspValThrLeuTyrGlnAlaPheAsn 60
 Db 156 CAGAACAATGAGATCAACCACTAGAAATGGAATGATGTTACTCTCTATAAAGCTTTCAAT 215
 QY 61 GlyLeuLeuAsnLysAsnSerGlnLeuLysThrLeuLeuAlaIleGlyGlyTpaAsnPhe 80
 Db 216 GACTTGAAAACAGGAACGAACTGAAACCCCTCTCGCAATGGAGCTGGAACTTT 275
 QY 81 GlyThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThr 100
 Db 276 GGAACCTGCTCTTCACTACCACTGGTTTCCACTCTCTCAGAACCGCCAGACCTTCATTACC 335
 QY 101 SerValIleIysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTpaGluTyr 120
 Db 336 TCAGTCATCAAAATTTCTGCGTCAGTATGGTTTGGAGCTGGACTGGGAATAC 395

QY 121 ProGlySerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValGlnGlu 140
 Db 396 CCAGGCTCAGCTGGGAGCCCTCTCAGGACAAGCATCTCTTCACTGTCTCTGGTGAAGGAA 455
 QY 141 MetArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThr 160
 Db 456 ATGCGTGAAGCTTTTGAGCAGGAGCTATTGAGCAACACAGCCAGACTGATGTTTACT 515
 QY 161 AlaAlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSer 180
 Db 516 GCTGCTGATAGCTGGTGGATTTCACATCCAGGCTGCTATGAGATCCCTGAATTTCT 575
 QY 181 GlnTyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr 200
 Db 576 AAGTACCTGGATTTTCATCATGTCATGATGATGATGATGATGATGATGATGATGATGAT 635
 QY 201 ThrGlyGluAsnSerProLeuTyrIleTyrProThrAspThrGlySerAsnAlaTyrLeu 220
 Db 636 ACTGGGAGAAATAGTCTCTTTTACAAATACCTACTAGACTGGTAGCAATGCCTACCTC 695
 QY 221 AsnValAspTyrValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeu 240
 Db 696 AATGTGGATTATGTCATGAACTATTGGAGAACATGAGCCCGCCAGCTGAGAACTCAT 755
 QY 241 ValGlyPheProThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIle 260
 Db 756 GTTGGATTTCCAGAGTATGGACACACCTTCCTCTGAGAACCCCTCTGATAATGGAATT 815
 QY 261 GlyAlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyIleTrp 280
 Db 816 GGTGCCCCCTACCTCTGTGTGGTGGCCCTGCTGGCCCTATACACAGA----- 860
 QY 281 AlaTyrTyrGluIleCysThrPheLeuLysAsnGlyAlaThrGlnGlyTrp 297
 Db 861 -----CAGGCTGG 869

RESULT 8
 BG867815 838 bp mRNA linear EST 29-MAY-2001
 LOCUS 602786336F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4912544 5',
 mRNA sequence.
 ACCESSION BG867815
 VERSION BG867815.1 GI:14218355
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 838)
 AUTHORS NIH-MGC http://imgc.nci.nih.gov/
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs@mail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM10816 row: g column: 09
 High quality sequence stop: 824.
 Location/Qualifiers
 1..838
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:4912544"
 /clone_lib="NCI CGAP_SG2"
 /lab_host="DH10B (T1 phage-resistant)"
 /note="Organ: salivary gland; Vector: pCMV-SPORT6; Site_1:

FEATURES
 source

NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo
dT. Average insert size 1.3 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT 209 a 211 c 198 g 220 t

ORIGIN

Alignment Scores:

Pred. No.: 1,47e-124 Length: 838
Score: 1270.00 Matches: 229
Percent Similarity: 93.18% Conservatives: 17
Best Local Similarity: 86.74% Mismatches: 18
Query Match: 50.68% Indels: 0
DB: 12 Gaps: 0

US-10-004-219B-14 (1-455) x BG867815 (1-838)

```
Qy 1 TyrGlnLeuThrCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPhe 20
Dy 34 TACAATCTGATATGCTATTTTCAACCACTGGGCCAGATATCGGCCAGGTCTGGGGAGCTTC 93
Qy 21 MetProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArg 40
Dy 94 AAGCTGTAGTAACTTAACCTTCCTGCTGTACTCACTGACTATGCTTTGCTGGGATG 153
Qy 41 GlnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrGlnAlaPheAsn 60
Dy 154 CAGAACATGATGATCACCACCATAGATGAATGATGTTACTTCTATAAAGCTTTCAAT 213
Qy 61 GlyLeuAsnLeuAsnSerGlnLeuIleThrLeuLeuAlaIleGlyGlyTrpAsnPhe 80
Dy 214 GACTTGAAACACGAGACAGCAACCTTCTGGAATTTGGAGGCTGGAACTTT 273
Qy 81 GlyThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThr 100
Dy 274 GGAATGCTCTCTTTCACATACATGTTTCCACTTCTCAGAACCCGACACCTTCATTACC 333
Qy 101 SerValIleLeuPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyr 120
Dy 334 TCAGTCATCAATTTCTGGCTCAGTATGGTTTGTAGTGCATGGACCTGGAGTGAATAC 393
Qy 121 ProGlySerArgGlySerProProGlnAspHisLeuPheThrValLeuValGlnGlu 140
Dy 394 CCAGGCTCAGCTGGGAGCCCTCTCAGGACAAGCATCTCTTCACTGCTGCTGGTGAAGAA 453
Qy 141 MetArgGluAlaPheGlnGluGlnAlaGlyGlnIleAsnLysProArgLeuMetValThr 160
Dy 454 ATGCTGAAGCTTTTGGAGGAGGCTTATGAGACCAACAGGCCAGACTGATGTTTACT 513
Qy 161 AlaAlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSer 180
Dy 514 GCTGCTGTAGCTGGTGGGATTTTCAACATCCAGGCTGGCTATGAGATCCCTGAATTTCT 573
Qy 181 GlnTyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTyr 200
Dy 574 AAGTACCTGGATTTTCCATCATCATATGATGATGATGATGATGATGATGATGATGATGAT 633
Qy 201 ThrGlyGluAsnSerProLeuTyrLysTyrProThrAspThrGlySerAsnAlaTyrLeu 220
Dy 634 ACTGGGGAGATGATGCTCTTTTACAAATACCTTACTGAGACTGGTAGCAATGCCCTACCTC 693
Qy 221 AsnValAspTyrValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeuIle 240
Dy 694 AATGTGGATTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 753
Qy 241 ValGlyPheProThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIle 260
Dy 754 GTTGATTTCCAGAGTTTGGACACACCTTCACTCTGAGAAACCCCTCTGATAATGATTT 813
Qy 261 GlyAlaProThr 264
Dy 814 GGTGCCCTACT 825
```

RESULT 9

BI517797
LOCUS

DEFINITION 603042049f1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5182357 5',
mRNA sequence.

ACCESSION

VERSION BI517797.1 GI:15342589

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (Bases 1 to 1038)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: scapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L14M1455 row: a column: 14
High quality sequence start: 3
High quality sequence stop: 788.

FEATURES

source

1..1038
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5182357"
/clone_lib="NIH_MGC_116"
/lab_host="DH10B"
/note="Organ: pooled colon, kidney, stomach; Vector:
pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
source anonymous pool of 3 colons, age 26 yo male, 49 yo
female, 71 yo male colon; 46 yo male kidney, and pool of 2
stomachs, 62 yo male and 70 yo female. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.4 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
023. Note: this is a NIH MGC Library."

BASE COUNT

ORIGIN

273 a 275 c 293 g 197 t

Alignment Scores:

Pred. No.: 2,17e-122 Length: 1038

Score: 1251.00 Matches: 237

Percent Similarity: 92.69% Conservatives: 4

Best Local Similarity: 91.15% Mismatches: 13

Query Match: 49.92% Indels: 6

DB: 13 Gaps: 2

US-10-004-219B-14 (1-455) x BI517797 (1-1038)

Qy 202 GlyGluAsnSerProLeuTyrLysTyrProThrAspThrGlySerAsnAlaTyrLeuAsn 221

Dy 2 GGAGAGAACAGCCCTCTACAAATACCGACTGACACCGGCAGCACGCTACCTCAAT 61

Qy 222 ValAspTyrValMetAsnTyrTrpLysAspAsnGlyAlaProAlaGluLysLeuIleVal 241

Dy 62 GTGATTATGTCATGAACACTACTGGAAGACAACTTCACTCTGACCAACCTCCACACTGGAAATGCT 121

Qy 242 GlyPheProThrTyrGlyHisAsnPheIleLeuSerAsnProSerAsnThrGlyIleGly 261

Dy 122 GGATTCCTACTATGACACAACTTCACTCTGACCAACCTCCACACTGGAAATGCT 181

Qy 262 AlaProThrSerGlyAlaGlyProAlaGlyProTyrAlaLysGluSerGlyIleTrpAla 281

Dy 182 GCCCCACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 241

LOCUS BG866417 809 bp mRNA linear EST 29-MAY-2001
DEFINITION 602785352F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4911421 5',
mRNA sequence.
ACCESSION BG866417
VERSION BG866417.1 GI:14216957
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 809)
NIH-MGC <http://mgc.nci.nih.gov/>.
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LHAM0813 row: h column: 14
High quality sequence stop: 804.
Location/Qualifiers
1. 809
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4911421"
/clone_lib="NCI CGAP SG2"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site 1:
NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo
dT. Average insert size 1.3 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 193 a 200 c 208 g 207 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 4,536-121 Length: 809
Score: 1237.00 Matches: 223
Percent Similarity: 90.71% Conservative: 21
Best Local Similarity: 82.90% Mismatches: 24
Query Match: 49.36% Indels: 1
DB: 12 Gaps: 0
US-10-004-219B-14 (1-455) x BG866417 (1-809)
QY 71 ThrLeuLeuAlaIleGlyClyTrpAsnPheGlyThrAlaProPheThrAlaMetValSer 90
DB 2 ACCCTCTGGCAATGGAGCGTGGAACTTGAACCTGCTCTTTCACCTACCATGTTTCC 61
QY 91 ThrProGluAsnArgGlnThrPheIleThrSerValIleLysPheLeuArgGlnTyrGlu 110
DB 62 ACTTCTCAGACCGCCAGACCTTCATTACCTCAGTCATCAATTTCTGGTCAGTATGGG 121
QY 111 PheAspGlyLeuAspPheAspTrpGluTyrProGlySerArgGlySerProProGlnAsp 130
DB 122 TTTGATGACTGGACTGGACTGGGAATACCCAGGCTCAGCTGGGAGCCCTCTCAGGAC 181
QY 131 LysHisLeuPheThrValLeuValGlnGluMetArgGluAlaPheGluGlnGluAlaLys 150
DB 182 AMGATCTCTTCTCTCTGCTGGTGAAGAAATGCGTGAAGCTTTTGAGCAGGAGCTATT 241
QY 151 GlnIleAsnLysProArgLeuMetValThrAlaAlaValAlaAlaGlyIleSerAsnIle 170
DB 242 GAGACACAGCCAGCCAGCTGATGTTACTGCTGTAGCTGTGGATTTCCACATC 301
QY 171 GlnSerGlyTyrGluIleProGlnLeuSerGlnTyrLeuAspTyrIleHisValMetThr 190
DB 302 CAGGCTGGCTATGAGATCCCTGAACCTTTCTAAGTACCTGGATTTTCATCATGTCATGACA 361

QY 191 TyrAspLeuHisGlySerTrpGluGlyTyrThrGlyGluAsnSerProLeuTyrIlystYr 210
DB 362 TATGACCTCCATGGCTCTCTGGAGGGCTACTCTGGGAGATAGTCTCTTTTACAAATAC 421
QY 211 ProThrAspThrGlySerAsnAlaTyrLeuAsnValAspTyrValMetAsnTyrTrpLys 230
DB 422 CCTACTGAGACTGGTAGCAATGCTTACCTCAATGGATTTATGTCATGAACCTATTGGAAG 481
QY 231 AspAsnGlyAlaProAlaGluLysLeuIleValGlyPheProThrTyrGlyHisAsnPhe 250
DB 482 AACATGGAGCCCGCCAGCTGAGAAGCTCATTTGTGATTTCCAGAGTATGGACACACCTTC 541
QY 251 IleLeuSerAsnProSerAsnThrGlyIleGlyAlaProThrSerGlyAlaGlyProAla 270
DB 542 ATCTGTGAGAAACCTCTGTATAATGGAATGGTGGCCCTACTCTGTGTGATGGCCCTGCT 601
QY 271 GlyProTyrAlaLysGluSerGlyIleTrpAlaTyrTyrGluIleCysThrPheLeuLys 290
DB 602 GGGCCCTATACACAGACAGCGCTGGGTTCTGGGCCTACTATGAGATNTGACACCTTTCTGAGA 661
QY 291 AsnGlyAlaThrGlnGlyTyrAspAlaProGlnGluValProTyrAlaTyrGlnGlyAs 310
DB 662 AGTGAGGCCACTGAGGTCTGGGATGCTCCCAAGAGTGGCCCTATGCTATAGGCCAA 721
QY 310 nValTrpValGlyTyrAspAsnIleLysSerPheAspIleLysAlaGlnTyrLeuLysHi 330
DB 722 CGAGTGGCTTGGCTATGACATTCACAGAGTTCAGTGTAAAGGGTCAGTGGCTTAAGCA 781
QY 330 sAsnLysPheGlyGlyAlaMetVal 338
DB 782 GAACAATTTGGAGTGCCATGATC 806
RESULT 12
BG8669304 764 bp mRNA linear EST 29-MAY-2001
LOCUS 602788819F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4920173 5',
mRNA sequence.
ACCESSION BG869304
VERSION BG869304.1 GI:14219844
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 764)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LHAM0836 row: e column: 06
High quality sequence stop: 760.
Location/Qualifiers
1. 764
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4920173"
/clone_lib="NCI CGAP SG2"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site 1:
NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo
dT. Average insert size 1.3 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 198 a 195 c 180 g 191 t

ORIGIN

Alignment Scores:
Pred. No.: 3,97e-118 Length: 764
Score: 1209.00 Matches: 219
Percent Similarity: 94.35% Conservative: 15
Best Local Similarity: 88.31% Mismatches: 14
Query Match: 48.24% Indels: 0
DB: 12 Gaps: 0

US-10-004-219B-14 (1-455) x BG869304 (1-764)

QY 8 ThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPheMetProAspAsnIleAspPro 27
DB 13 ACCAATGGGCCCAGTATGCCAGGCTTCAGGAGCTTCAGGCTGATGACATTAACCCC 72
QY 28 CysLeuYerThrsIleuLeuTyrAlaPheAlaGlyArgGlnAsnGlnIleThrThr 47
DB 73 TGCCTGTGTACTCACCTGATCTATGCTTTTGTGGGATGCAGAACATGAGATCACCC 132
QY 48 IleGluTrpAsnAspValThrLeuTyrGlnAlaPheAsnGlyLeuLysAsnSer 67
DB 133 ATAGATGAATGATGATCTCTATAGCTTTCAATGACTTGAAGAACAGGACAGC 192
QY 68 GlnLeuLysThrLeuLeuAlaIleGlyGlyTyrAsnPheGlyThrAlaProPheThrAla 87
DB 193 AAACCTGAAACCCCTCTGGCAATGGAGCTGGAACTTTGGAACTGCTCTTTCACTACC 252
QY 88 MetValSerThrProGluAsnArgGlnThrPheIleThrSerValIleLysPheLeuArg 107
DB 253 ATGGTTTCCACTTCTCAGAACCCGACAGCTTCATTAACCTCACTCATCAATTTCTGCT 312
QY 108 GlnTyrGluPheAspGlyLeuAspPheAspTyrGluTyrProGlySerArgGlySerPro 127
DB 313 CAGTATGGGTGTGATGGACTGGACCTGGAGTGGGAATACCCAGGCTCAGCTGGAGCCCT 372
QY 128 ProGlnAspLysHisLeuPheThrValLeuValGlnGluMetArgGluAlaPheGluGln 147
DB 373 CCTCAGACACAGCATCTCTCACTGCTCTGCTGGTGAAGAAATCGTGAAGCTTTTGAGCAG 432
QY 148 GluAlaLysGlnIleAsnLysProArgLeuMetValThrAlaAlaValAlaGlyIle 167
DB 433 GAGGCTATTGAGAGCAACAGGCCACAGCTGATGTGTACTGCTGTAGCTGTGGGATT 492
QY 168 SerAsnIleGlnSerGlyTyrGluIleProGlnLeuSerGlnTyrLeuAspTyrIleHis 187
DB 493 TCAACATCCAGCTGGCTATGAGATCCCTGAACCTTTCTAAGTACCTGGATTTCATCCAT 552
QY 188 ValMetThrTyrAspLeuHisGlySerTrpGluGlyTyrThrGlyGluAsnSerProLeu 207
DB 553 GTCATGACATATGACCTCCATGCTCTCTGGAGGCTACACTGGGAGAAATAGTCTCTT 612
QY 208 TyrLysTyrProThrAspThrGlySerAsnAlaTyrLeuAsnValAspTyrValMetAsn 227
DB 613 TACAAATACCTACTAGAGCTGGTAGCAATGCCCTACCTCAATGTGATTTATGTGATGAAC 672
QY 228 TyrTrpLysAspAsnGlyAlaProAlaGluLysLeuIleValGlyPheProThrTyrGly 247
DB 673 TATTGAGAGCAATGAGCCCCAGCTGAGAGCTCATTTGTTGGATTCCAGATGATGGA 732
QY 248 HisAsnPheIleLeuSerAsnPro 255
DB 733 CACACCTTTCATCTGAGAAACCT 756

RESULT 13
BG872647
LOCUS 763 bp mRNA linear EST 29-MAY-2001
DEFINITION 602793705F1.NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4924897 5',
mRNA sequence.
ACCESSION BG872647
VERSION BG872647.1 GI:14223187
KEYWORDS EST.
SOURCE house mouse.

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 763)

NIH-MGC <http://mgc.nci.nih.gov/>

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLN10848 row: j column: 02

High quality sequence stop: 760.

FEATURES

source

1..763

/organism="Mus musculus"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone="IMAGE:4924897"

/lab_host="NCI CGAP_SG2"

/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site 1:

NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo

dT. Average insert size 1.3 kb. Constructed by Life

Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT 199 a 190 c 178 g 196 t

ORIGIN

Alignment Scores:
Pred. No.: 6,47e-118 Length: 763
Score: 1207.00 Matches: 222
Percent Similarity: 93.68% Conservative: 15
Best Local Similarity: 87.75% Mismatches: 16
Query Match: 48.16% Indels: 1
DB: 12 Gaps: 0

US-10-004-219B-14 (1-455) x BG872647 (1-763)

QY 3 LeuThrCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPheMetPro 22
DB 1 CTGATATGCTATTTTCAACCACTGGGCCAGTATCGCCAGGCTCGGGAGCTTCAAGCCT 60
QY 23 AspAsnIleAspProCysLeuCysThrHisIleuIleTyrAlaPheAlaGlyArgGlnAsn 42
DB 61 GATGACATTAACCCCTGCTGTACTCACCTGATCTATGCCCTTTGCTGGGATGCAGAAC 120
QY 43 AsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrGlnAlaPheAsnGlyLeu 62
DB 121 AATGAGATCACCAACATAGATGGAATGATGTTACTCTCTATAAGCTTTCAATGACTTG 180
QY 63 LysAsnLysAsnSerGlnLeuLysThrLeuLeuAlaIleGlyGlyTyrAsnPheGlyThr 82
DB 181 AAAACAGACAGACAGCAACTGAAACCTCTCGCAATTGGAGGCTGGAACCTTTGGAAC 240
QY 83 AlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThrSerVal 102
DB 241 GCTCCTTTCACTTACCATTGCTTCCACTTCTCAGAACCCGACAGACCTTCATTACCTCAGTC 300
QY 103 IleLysPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyrProGly 122
DB 301 ATCAAAATTTCTGCGTCAGTATGGGTTTGTATGACTGGACCTGGAGTGGAAATACCAGC 360
QY 123 SerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValGlnGluMetArg 142
DB 361 TCACGTGGGAGCCCTCTCTCAGACAGCATCTCTTCACTGTCTCTGGTGAAGAAATGCT 420
QY 143 GluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThrAlaAla 162

Db 421 GAAGCTTTTGAGCAGGAGCTATTGAGAGCAACAGCCGCCAGACTGATGGTTACTGCTGCT 480
 Qy 163 ValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSerGlnTyr 182
 Db 481 GTAGCTGTGGGATTTCCAAATCCAGCTGCTGATGAGATCCCTGAACTTCTTAAGTAC 540
 Qy 183 LeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTyrGluGlyTyrThrGly 202
 Db 541 CTGGATTTCATCCATGTCATGACATATGACCTCCATGCTCTCTGGAGGGCTACACTGGG 600
 Qy 203 GluAsnSerProLeuTyrLysTyrProThrAspThrGlySerAsnAlaTyrLeuAsnVal 222
 Db 601 GAGAACTGCTCTTTTACAAATACCTTACTGAGACTGGTAGCAATGGCTTACCTCAATGTG 660
 Qy 223 AspTyrValMetAsnTyrTyrLysAspAsnGlyAlaProAlaGluLysLeuIleValGly 242
 Db 661 GATTATGTCATGAACTATTGGNAGAACATGGAGC-CCAGCTGAGAGCTCATTTGTGGA 719
 Qy 243 PheProThrTyrGlyHisAsnPheIleLeuSerAsnPro 255
 Db 720 TTCCAGAGTATGAGACACACCTTTCATCTCTGAGAAACCTT 758

RESULT 14
 BG868741
 LOCUS 602787570F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4913800 5',
 DEFINITION mRNA sequence.
 ACCESSION BG868741
 VERSION BG868741.1 GI:14219281
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (Bases 1 to 837)
 TITLE NIH-MGC http://mgi.nci.nih.gov/.
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/ILNL at:
 http://image.llnl.gov
 Plate: LHAM10819 row: k column: 17
 High quality sequence stop: 836.

FEATURES
 source
 1..837
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:4913800"
 /clone_lib="NCI CGAP SG2"
 /lab_host="DH10B (T1 phage-resistant)"
 /note="Organ: salivary gland; Vector: pCMV-SPORT6; Site 1:
 NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo
 dt. Average insert size 1.3 kb. Constructed by Life
 Technologies. Note: this is a NCI CGAP Library."
 BASE COUNT 212 a 211 c 197 g 217 t

ORIGIN
 Alignment Scores:
 Pred. No.: 1.22e-117 Length: 837
 Score: 1205.00 Matches: 219
 Percent Similarity: 92.58% Conservative: 18
 Best Local Similarity: 85.55% Mismatches: 18
 Query Match: 48.08% Indels: 1
 DB: 12 Gaps: 0
 US-10-004-219B-14 (1-455) x BG868741 (1-837)

Qy 1 TyrGlnLeuThrCysTyrPheThrAsnTyrAlaGlnTyrArgProGlyLeuGlyArgPhe 20
 Db 66 TACAATCTGATATGCTATTTTCCACCACTGGGCCAGTATCGCCAGGTCCTGGGAGCTTC 125
 Qy 21 MetProAspAsnIleLeuAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArg 40
 Db 126 AAGCCTGATGACATTAACCCCTGCTGTGTACTCACCTGATCTATGCTCTTGTGGGATG 185
 Qy 41 GlnAsnGluIleThrThrIleGluTyrAsnAspValThrLeuTyrGlnAlaPheAsn 60
 Db 186 CAGAACATGAGATCACCACTAGATGGAATGGAATGATTTACTCTCTATPAAGCTTTTCAAT 245
 Qy 61 GlyLeuLysAsnLysAsnSerGlnLeuLysThrLeuLeuAlaIleGlyGlyTyrAsnPhe 80
 Db 246 GACTTGAACACAGACAGCAAACTGAAACCCCTCTCTGGCAATTTGGAGGCTGGAACCTT 305
 Qy 81 GlyThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThr 100
 Db 306 GGAATCTGCTCTTTCACCTACCATGCTTTCCACTTCTCAGAACCCGACAGCTTCAATACC 365
 Qy 101 SerValIleLysPheLeuAtgGlnTyrCluPheAspGlyLeuAspPheAspTyrGluTyr 120
 Db 366 TCAGTCATCAAAATTTCTGCGTCAGTATGGGTTTGTAGGACTGGACCTGGGAAATAC 425
 Qy 121 ProGlySerArgGlySerProProGlnAspLysHisLeuPheThrValLeuValGlnGlu 140
 Db 426 CCAGGCTCACCTGGGAGCCCTCTCCAGGACCAAGCATCTCTTCACTGTCTGGTGAAGGA 485
 Qy 141 MetArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThr 160
 Db 486 ATGCTGGAAGCTTTTGGAGCAGGAGCTATTGAGAGCAACAGGCCCCAGACTGATGTTACT 545
 Qy 161 AlaAlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSer 180
 Db 546 GCTGCTGTAGCTGTGGGATTTTCCAAATCCAGGCTGCTGATGAGATCCCTGAACTTTCT 605
 Qy 181 GlnTyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTyrGluGlyTyr 200
 Db 606 AAGTACCTGGATTTTCATCCATGTCATGACATATGACCTCCATGGCTCTGGGAGGGCTAC 665
 Qy 201 ThrGlyGluAsnSerProLeuTyrLysTyrProThrAspThrGlySerAsnAlaTyrLeu 220
 Db 666 ACTGGGAGATAGTCTCTTTTCAAAATCCCTACTGAGACTGGTAGCAATGCCCTTACCTC 725
 Qy 221 AsnValAspTyrValMetAsnTyrTyrLysAspAsnGlyAlaProAlaGluLysLeuIle 240
 Db 726 AATGTGGATTATGTCATGAACCTATTGGGAGCAACAATGGAGCCCGCCAGTGAAGCTCAT 785
 Qy 241 ValGlyPheProThrTyrGlyHisAsnPheIleLeuSerAsnPro 255
 Db 786 GTTGGATTCCAGAGTATGGACACACCTTCAATCTCTGAGAAACCTT 831

RESULT 15
 BQ219632
 LOCUS 939 bp mRNA linear EST 02-MAY-2002
 DEFINITION AGSCOURT 7572993 NCI CGAP_St1 Mus musculus cDNA clone
 IMAGE:6051913 5', mRNA sequence.
 ACCESSION BQ219632
 VERSION BQ219632.1 GI:20401032
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (Bases 1 to 939)
 TITLE NIH-MGC http://mgi.nci.nih.gov/.
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM13306 row: i column: 02
High quality sequence stop: 654.
Location/Qualifiers
1. .939
/organism="Mus musculus"
/db_xrefs="taxon:10090"
/clone_image="6051913"
/lab_host="DH10B (T1-resistant)"
/note="Organ: stomach; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.77 kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 232 a 246 c 224 g 237 t
ORIGIN

Alignment Scores:
Pred. No.: 3,06e-117 Length: 939
Score: 1202.00 Matches: 233
Percent Similarity: 87.41% Conservative: 17
Best Local Similarity: 81.47% Mismatches: 29
Query Match: 47.96% Indels: 7
DB: 14 Gaps: 4

US-10-004-219B-14 (1-455) x BQ219632 (1-939)

```
QY 1 TyrGlnLeuThrCysTyrPheThrAsnTrpAlaGlnTyrArgProGlyLeuGlyArgPhe 20
DB 79 TACATCTGATGCTATTTTCCACCACTGGGCCAGTATCGGCCAGGTCTGGGAGCTTC 138
QY 21 MetProAspAsnIleAspProCysLeuCysThrHisLeuIleTyrAlaPheAlaGlyArg 40
DB 139 AAGCCTGATGACATTAAACCCCTGCTGTGTACTCACCTGATCTATGCTTGTCTGGATG 198
QY 41 GlnAsnAsnGluIleThrThrIleGluTrpAsnAspValThrLeuTyrGlnAlaPheAsn 60
DB 199 CAGAACATGAGATCAACCACTAGATGAATGATGTTACTCTCTATAAAGCTTTCAAT 258
QY 61 GlyLeuLeuAsnIleAsnSerGlnLeuLeuThrLeuAlaIleGlyTyrAsnPhe 80
DB 259 GACTTGAAAAACAGGACAGCAACCTCTCTGCAATTCGAGGCTGGAACTTT 318
QY 81 ThrAlaProPheThrAlaMetValSerThrProGluAsnArgGlnThrPheIleThr 100
DB 319 GGAACCTGCTCTTCACTACCATGGTTCCACTTCTCAGAACCGCCAGACCTTCATTACC 378
QY 101 SerValIleLeuPheLeuArgGlnTyrGluPheAspGlyLeuAspPheAspTrpGluTyr 120
DB 379 TCAGTCATCAAAATTTCTGCGTCAGTATGGTTGATGGACTGGACCTGGGAATAC 438
QY 121 ProGlySerArgGlySerProGlnAspLysHisLeuPheThrValLeuValGlnGlu 140
DB 439 CCAGGCTCAGTGGGAGCCCTCTCAGGACAGCATCTCTTCACTGCTCTGTTGAAGGAA 498
QY 141 MetArgGluAlaPheGluGlnGluAlaLysGlnIleAsnLysProArgLeuMetValThr 160
DB 499 ATGCGTGAAGCTTTTGAGCAGGAGGCTATTGAGAGCACAGGCCAGACTGATGGTTACT 558
QY 161 AlaAlaValAlaAlaGlyIleSerAsnIleGlnSerGlyTyrGluIleProGlnLeuSer 180
DB 559 GCTGCTGTAGCTGGTGGGATTTCCAAACATCCAGGCTGGCTATGAGATCCCTGAACCTTCC 618
QY 181 -GlnTyrLeuAspTyrIleHisValMetThrTyrAspLeuHisGlySerTrpGluGlyTy 200
DB 619 TAAGTACCTGGATTTCCATCCATGTCATGACATATGACCTCCATGGCTCTGGGAGGCTA 678
QY 200 rThrGlyGluAsnSerProLeuTyrIleTyrProThrAspThrGlySerAsnAlaTyrIle 220
```

```
DB 679 CCTGGGGAGATAGTCTCTTTACAAATACCTACTGAGACTGGTAGCAATGCCTACT 738
QY 220 uAsnValAspTyrValMetAsnTyrTrpLysAspAsnGlyAla-ProAlaGluLysLeuI 240
DB 739 CAATGTGGGATATGTCATGAACCTATTGGAAAAACATGTGAGGCCCCAGCTGAGAAGCTCA 798
QY 240 leValGlyPheProThrTyrGly---HisAsnPheIleLeuSerAsnProSerAsnThrG 259
DB 799 TTGTTGGATTCCCAAAATATGGAACACACACCTTCATCTCTGAAAAAACCTTCTTAATAATG 858
QY 259 ly-IleGlyAlaPro---ThrSerGlyAlaGlyPro---AlaGlyPro---TyrAlaLys 275
DB 859 GAAATTGGGGGGCCCCCTTACCTCTGGAGGATGGCCCCCTGTGGGGCCCCCAATACCCAAAA 918
QY 276 GluSerGlyIle 279
DB 919 AGGCGTTGGGTT 930
```

Search completed: June 30, 2003, 01:08:01.
Job time : 1512.62 secs